

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C07H 15/203, 15/04, 23/00, A61K 31/70</b>		<b>A2</b>	<b>(11) International Publication Number:</b> <b>WO 95/14028</b>
			<b>(43) International Publication Date:</b> 26 May 1995 (26.05.95)
<b>(21) International Application Number:</b> PCT/US94/13455			<b>45 Uppsala (SE).</b> LEPISTO, Matti [SE/SE]; Flygelvagen 257, S-224 73 Lund (SE). JONES, Charles, Hal [US/US]; 1104A Moorlands Drive, Saint Louis, MO 63110 (US). KIHLLBERG, Jan [SE/SE]; Havrevagen 16, S-240 10 Dalby (SE).
<b>(22) International Filing Date:</b> 18 November 1994 (18.11.94)			
<b>(30) Priority Data:</b> 08/154,035 18 November 1993 (18.11.93) US			
<b>(60) Parent Application or Grant</b> <b>(63) Related by Continuation</b> US 08/154,035 (CIP) Filed on 18 November 1993 (18.11.93)			
<b>(71) Applicants (for all designated States except US):</b> WASHINGTON UNIVERSITY [US/US]; Box 8013, 724 South Euclide Avenue, Saint Louis, MO 63110 (US). SYMBICOM AK-TIEBOLAG [SE/SE]; Tvistevagen 48, S-907 36 Umea (SE).			
<b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> HULTGREN, Scott [US/US]; 1637 Country Hill Lane, Balwin, MO 63021 (US). KUEHN, Meta [US/US]; 605 Clara Avenue, Saint Louis, MO 63112 (US). XU, Zheng [US/US]; 3201 January Avenue, Saint Louis, MO 63139 (US). OGG, Derek [GB/SE]; Artillerigaten 16B, S-752 37 Uppsala (SE). HARRIS, Mark [GB/SE]; Norbykallvagen 2, S-756			<b>(74) Agent:</b> COOPER, Iver, P.; Browdy and Neimark, Suite 300, 419 7th Street, N.W., Washington, DC 20004 (US).
			<b>(81) Designated States:</b> AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, HU, JP, KP, KR, KZ, LK, LU, LV, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).
<b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>			
<b>(54) Title:</b> COMPOUNDS AND PHARMACEUTICAL COMPOSITIONS FOR THE TREATMENT AND PROPHYLAXIS OF BACTERIAL INFECTIONS			
<b>(57) Abstract</b>  Novel methods for the treatment and/or prophylaxis of diseases caused by tissue-adhering bacteria are disclosed. By interacting with periplasmic molecular chaperones it is achieved that the assembly of pili is prevented or inhibited and thereby the infectivity of the bacteria is diminished. Also disclosed are methods for screening for drugs as well as methods for the de novo design of such drugs, methods which rely on novel computer drug modelling methods involving an approximative calculation of binding free energy between macromolecules. Finally, novel pyranosides which are believed to be capable of interacting with periplasmic molecular chaperones are also disclosed.			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

COMPOUNDS AND PHARMACEUTICAL COMPOSITIONS FOR THE TREATMENT  
AND PROPHYLAXIS OF BACTERIAL INFECTIONS

This invention was made with US government support under grant number R01AI29549 (S.J.H.) and training grant AI07172, both awarded by NIH. The US government has certain rights in the invention.

FIELD OF THE INVENTION

The present invention relates to methods for the treatment and/or prophylaxis of diseases caused by tissue-adhering pilus-forming bacteria by interaction with the binding between pilus subunits and periplasmic chaperones. The invention further relates to methods for identifying/designing substances capable of interacting with periplasmic chaperones and methods for identifying binding sites in periplasmic chaperones. Finally, the invention relates to novel substances capable of interacting with periplasmic chaperones as well as pharmaceutical preparations comprising substances capable of interacting with periplasmic chaperones.

BACKGROUND OF THE INVENTION

Pathogenic Gram-negative bacteria cause a number of pathological conditions such as bacteraemia, bacteria-related diarrhoea, meningitis and (very commonly) urinary tract infections, i.a. pyelonephritis, cystitis, urethritis etc.

Urinary tract infections are one of the major causes of morbidity in females. Despite the overall importance of urinary tract infections in women, there have been few efforts to apply novel strategies in order to treat and/or prevent these diseases. Commonly, conventional antibiotics are used to treat these infections, such as treatment with penicillins, cephalosporins, aminoglycosides, sulfonamides and tetracyclines; in the special case of urinary tract infections, urinary antiseptics such as nitrofurantoin and

nalidixic acid are employed, too. However, emerging anti-biotic resistance will in the future hamper the ability to successfully treat urinary tract infections. Multiple anti-biotic resistance among these uropathogens is increasing. It has been estimated that the annual cost for evaluation and treatment of women with urinary tract infections exceeds one billion dollars. In addition, approximately one-fourth of the yearly 4 billion dollar cost attributed to nosocomial infections is a consequence of urinary tract infections. Among the causative agents of urinary tract infections, *Escherichia coli* clearly predominates among Gram-negative bacteria.

Pathogenic gram negative bacteria, notably *Escherichia coli*, *Haemophilus influenzae*, *Salmonella enteritidis*, *Salmonella typhimurium*, *Bordetella pertussis*, *Yersinia pestis*, *Yersinia enterocolitica*, *Helicobacter pylori*, and *Klebsiella pneumoniae* owe part of their infectability to their ability to adhere to various epithelial tissues. Thus, e.g. *E. coli* adhere to the epithelial cells in the upper urinary tract in spite of the flushing effect of unidirectional flow of urine from the kidneys.

As indicated above, the above mentioned bacteria are involved in a variety of diseases: Urinary tract infections (*E. coli*), acute diarrhoea (*E. coli*, *Y. enterocolitica* and *Salmonella* spp), meningitis (*E. coli* and *H. influenzae*), whooping cough (*B. pertussis*), plague (*Y. pestis*), pneumonia and other respiratory tract infections (*K. pneumoniae*, *H. influenzae*) and peptic ulcer (*H. pylori*).

The initiation and persistence of many bacterial infections such as those described above is thought to require the presentation of adhesins on the surface of the microbe in accessible configurations which promote binding events that dictate whether extracellular colonization, internalization or other cellular responses will occur. Adhesins are often components of the long, thin, filamentous, heteropolymeric protein appendages known as pili, fimbriae, or fibrillae



(these three terms will be used interchangeably herein). The bacterial attachment event is often the result of a stereochemical fit between an adhesin frequently located at the pilus tip and specific receptor architectures on host cells, often comprising carbohydrate structures in membrane associated glycoconjugates.

Uropathogenic strains of *E. coli* express P and type 1 pili that bind to receptors present in uroepithelial cells. The adhesin present at the tip of the P pilus, PapG (pilus associated polypeptide G), binds to the Gal $\alpha$ (1-4)Gal moiety present in the globoseries of glycolipids, while the type 1 adhesin, FimH, binds D-mannose present in glycolipids and glycoproteins. Adhesive P pili are virulence determinants associated with pyelonephritic strains of *E. coli* whereas type 1 pili appear to be more common in *E. coli* causing cystitis. At least eleven genes are involved in the biosynthesis and expression of functional P pili; the DNA sequence of the entire pap gene cluster has been determined. P pili are composite heteropolymeric fibres consisting of flexible adhesive fibrillae joined end to end to pilus rods. The pilus rod is composed of repeating PapA protein subunits arranged in a right handed helical cylinder. Tip fibrillae which extend from the distal ends of each pilus rod were found to be composed mostly of repeating subunits of PapE arranged in an open helical conformation. The PapG adhesin was localized to the distal ends of the tip fibrillae, a location which is assumed to maximize its ability to recognize glycolipid receptors on eukaryotic cells. Two minor pilus components, PapF and PapK, are specialized adaptor proteins found in the tip fibrillum. PapF links the adhesin moiety to the fibrillum while PapK joins the fibrillum to the pilus rod. The composite architecture of the P pilus fibre reveals the strategy used by uropathogenic *E. coli* to present the PapG adhesin to eukaryotic receptors. The rigid PapA rod extends the adhesin away from interference caused by LPS and other components at the bacterial cell surface while the flexible fibrillum allows PapG steric freedom to recognize

and bind to the digalactoside moiety on the uroepithelium. With a few exceptions, the structural organization of type 1 pili is very similar to that described for P pili. In type 1 pili, the mannose binding fibrillar tip adhesin is known as  
5 *FimH*.

The assembly of virulence-associated pili in Gram-negative pathogens requires the function of periplasmic chaperones. Molecular chaperones are vital components of all living cells, prokaryotic and eukaryotic. Chaperones serve a variety  
10 of cellular functions including folding, import and export of proteins in various cellular compartments (Gething and Sambrook, 1992). Thus, a periplasmic chaperone is a molecular chaperone which exerts its action in the periplasmic space in bacteria.

15 *PapD* and *FimC* are the periplasmic chaperones that mediate the assembly of P and type 1 pili, respectively. Detailed structural analyses have revealed that *PapD* is the prototype member of a conserved family of periplasmic chaperones in Gram-negative bacteria. These chaperones have a function  
20 which is part of a general strategy used by bacteria to cap and partition interactive subunits imported into the periplasmic space into assembly competent complexes, making non-productive interactions unfavourable. Determination of the three-dimensional structure of *PapD* revealed that it consists  
25 of two immunoglobulin-like domains oriented in a boomerang shape such that a cleft is formed. *PapD* binds to each of the pilus subunit types as they emerge from the cytoplasmic membrane and escorts them in assembly-competent, native-like conformations from the cytoplasmic membrane to outer membrane  
30 assembly sites comprised of *PapC*. *PapC* has been termed a molecular usher since it receives chaperone-subunit complexes and incorporates, or ushers, the subunits from the chaperone complex into the growing pilus in a defined order.

With the exception of the type IV class of pili, all other  
35 genetically well characterized pilus systems in Gram-negative

prokaryotes contain a gene analogous to *PapD* (Normark et al., 1986; Hultgren et al. 1991); cf. also table A. *FanE*, *faeE*, *sfaE*, *ClpE* and *f17-D* have been sequenced (Lintermans, 1990; Schmoll et al., 1990; Bertin Y et al., 1993; Bakker et al., 1991) and encode pilus chaperones required for the assembly of K99, K88, S and F17 pili, respectively, in *E. coli*. The assembly of *Klebsiella pneumoniae* type 3 pili and *Haemophilus influenzae* type b pili requires the *mrkb* and *hifB* gene products, respectively (Gerlach et al., 1989,; Allen et al., 1991). The structure-function relationships of all of these chaperones have been analyzed using their amino acid sequences and information from the crystal structure of *PapD* (Anders Holmgren et al., 1992). The results have provided insight into the molecular intricacies that have been evolutionarily conserved in this class of proteins and suggested significant structural similarities to immunoglobulins.

*PapD* is thus the prototype member of a family of periplasmic chaperone proteins which are necessary for the correct supramolecular assembly of bacterial pili. Chaperones such as *PapD* in *E. coli* are required to bind the above-mentioned pilus proteins imported into the periplasmic space, partition them into assembly competent complexes and prevent non-productive aggregation of the subunits in the periplasm (Dodson et al., 1993).

In the absence of an interaction with the chaperone, pilus subunits aggregate and are proteolytically degraded (Kuehn, Normark, and Hultgren, 1991). It has recently been discovered (Strauch, Johnson and Beckwith, 1989) that the DegP protease is greatly responsible for the degradation of pilin subunits in the absence of the chaperone. This discovery has allowed the elucidation of the fate of pilus subunits expressed in the presence or absence of the chaperone using monospecific antisera in western blots of cytoplasmic membrane, outer membrane and periplasmic proteins prepared according to standard procedures. Expression of *papG* or *papA* in the *degP41* strain (a DegP<sup>-</sup> *E. coli* strain) in the absence of a chaperone

was toxic to the bacteria due to the accumulation of these proteins in the cytoplasmic membrane suggesting that the chaperone was required for subunit import into the periplasmic space. The severity of the growth defect was related to the level of expression of the pilin subunit, and was generally more dramatic with *PapG* than *PapA*. Co-expression of *PapD* under the control of the inducible arabinose promoter rescued the growth defect associated with subunit expression in the *degP41* strain and allowed *PapG* to be imported into the periplasm.

Up to this time, little has elucidated about the molecular recognition motifs of chaperones. Cytoplasmic chaperones such as *SecB*, *GroEL* and *DnaK* bind to a diverse group of unfolded target proteins in a sequence independent manner (Gething and Sambrook, 1992). Recently, it has been suggested that *DnaK* binds mainly via the target's peptide backbone (Landry et al., 1992) and that *GroEL* may rely more on side-chain hydrophobicity and the ability of the target sequence to form an amphipathic  $\alpha$ -helix (Landry and Gierasch, 1991). *PapD* differs, however, in that it seems to bind its target proteins in folded conformations (Kuehn et al., 1991).

The three-dimensional structure of *PapD* has previously been solved (Holmgren and Brändén, 1989). This has shown *PapD* to consist of two globular domains positioned such that the overall shape of the molecule resembles a boomerang with a cleft between the two domains. Each domain is a  $\beta$ -barrel structure formed by two antiparallel  $\beta$ -pleated sheets, packed tightly together to form a hydrophobic core, with a topology similar to that of an immunoglobulin fold. The N-terminal domain of *PapD* most resembles an Ig variable domain whilst the C-terminal domain of *PapD* resembles CD4 (Wang et al., 1990, Ryu et al., 1990) and the human growth hormone receptor (de Vos et al., 1992).

A structural alignment between *PapD* and several periplasmic chaperones predicted to have a similar immunoglobulin-like

structure has identified invariant, highly conserved and variable residues within this protein family (Holmgren et al., 1992). Most conserved residues seem to participate in maintaining the overall structure and orientation of the domains towards one another. However, two conserved residues, Arg-8 and Lys-112 are surface exposed and oriented towards the cleft between the domains. Site-directed mutagenesis of the Arg-8 amino acid has shown that it form at least part of the pilus subunit binding pocket (Holmgren et al., 1992; Kuehn et al., 1993).

From a sequence analysis of a number of the above-mentioned pilus subunit proteins, it has been observed that they possess a number of common features including homologies at the C termini (see also example 2). It is thought that these similarities in sequence may be responsible for some function common to all the pilus proteins, such as binding to their periplasmic chaperone. Indeed, the C-terminal region of the P-pili adhesin *PapG* has already been shown to be important in the *in vivo* binding to *PapD* (Hultgren et al., 1989). Table A lists 16 periplasmic proteins, all involved in assembly of cell surface structures in pathogenic bacteria and all with significant homology with *PapD*.

Table A

	Organism	Chaperone Structure Assembled	Reference	
	<i>E. coli</i>	PapD	P pili	Lund et al., 1987
	<i>E. coli</i> *	FimC	Type 1 pili	Klemm et al., 1992
5	<i>E. coli</i>	Sfae	S pili	Schmoll et al., 1990
	<i>E. coli</i>	FaeE	K88 pili	Bakker et al., 1991
	<i>E. coli</i>	FaeE	K99 pili	Bakker et al., 1991
	<i>E. coli</i>	CS3-1	CS3 pili	Jalajakumari et al., 1989
	<i>E. coli</i>	F17D	F17 pili	Lintermans et al., 1990
10	<i>E. coli</i>	ClpE	CS31 A	Bertin et al., 1993
	<i>E. coli</i>	EcpD	not identified	Raina et al., 1993
	<i>E. coli</i>	CssC	Antigen CS6	
	<i>E. coli</i>	NfaE	Nonfimbrial adhesin 1	
	<i>E. coli</i>	AggD	Aggregative Adherence Fimbria 1	
15	<i>K. pneumoniae</i>	MrkB	Type 3 pili	Allen et al., 1991
	<i>B. pertussis</i>	FimB	Type 2 & 3 pili	Locht et al., 1992; Willems et al. 1992
	<i>S. enteritidis</i>	SefB	unknown pili	Clouthier et al., 1993
	<i>S. typhimurium</i>	PefD	PEF	
	<i>H. influenzae</i>	HifB	unknown pili	Smith et al., 1993
20	<i>Y. enterocolitica</i>	MyfB	Myf fibrillae	Iriarte et al., 1993
	<i>Y. pestis</i>	PsAB	pH6 antigen	Lindler et al., 1992
	<i>Y. pestis</i>	CaflM	F1 envelope antigen	Galyov et al., 1991
	<i>P. mirabilis</i>	MrpD	MR/P Fimbriae	
	?	YehC	?	
25	* This chaperone is present in all Enterobacteriaceae, since all members of this family produce type 1 pili.			

In summary, the three-dimensional structure of PapD as well as the function of PapD and other periplasmic chaperones are known, whereas the exact motif of binding between PapD and the pilus subunits has been unknown until now.

#### DISCLOSURE OF THE INVENTION

According to the present invention, it has been realized that the above-mentioned characteristics of PapD and other related

chaperones make them interesting targets for drugs designed primarily to reduce the pathogenicity of bacteria which adhere by means of pili; for example, a drug which blocks the binding between a chaperone and the pilus-subunits (thereby  
5 interfering with the assembly of the intact pilus) will interfere with the formation of intact pili, thereby reducing bacterial capacity to adhere to host epithelium.

In order to design such a drug it is of great value that the motif of binding between the chaperone and the pilus pro-  
10 tein(s) is known in detail, in order to develop the method to effectively identify compounds capable of blocking this binding.

An aspect of the current invention relates to a method for the treatment and/or prophylaxis of diseases caused by tis-  
15 sue-adhering pilus-forming bacteria, comprising preventing, inhibiting or enhancing binding between at least one type of pilus subunit and at least one type of molecular chaperone in the pilus-forming bacteria, which molecular chaperone binds pilus subunits during transport of these pilus subunits  
20 through the periplasmic space and/or during the process of assembly of the intact pilus.

As used herein, the term "pilus", "fimbria", or "fibrilla", relates to fibrillar heteropolymeric structures embedded in the outer membrane of many tissue-adhering pathogenic  
25 bacteria, notably the pathogenic gram negative bacteria. In the present specification the terms pilus, fibrillum and fimbria will be used interchangeably. A pilus is, as explained above, composed of a number of "pilus subunits", which constitute distinct functional parts of the intact  
30 pilus. A very important pilus subunit is the "adhesin", the pilus subunit which is responsible for the tissue-binding capacity of the bacterium.

By the term "molecular chaperone" is meant a molecule which in living cells has the responsibility of binding to peptides

in order to mature the peptides in a number of ways. Many molecular chaperones are involved in the process of folding of peptides into their native conformation whereas other molecular chaperones are involved in the process of export  
5 out of or import into the cell of peptides. Specialized molecular chaperones are "periplasmic chaperones", which are bacterial molecular chaperones exerting their main actions in the "periplasmic space" (the space between the inner and outer bacterial membrane). Periplasmic chaperones are  
10 involved in the process of correct assembly of intact pili. When used herein, the simple term "chaperone" designates a molecular, periplasmic chaperone if nothing else is indicated.

When using the phrase "one type of" is meant that the pilus  
15 subunit or the chaperone in question is of one distinct species. However, especially the fact that there is extensive homology between different species of periplasmic molecular chaperones renders it likely that the interference with one type of chaperone using e.g. a compound will make it possible  
20 to also use the compound in interference with other chaperones.

The phrase "preventing, inhibiting or enhancing binding between pilus subunits and at least one molecular chaperone in the pilus-forming bacteria" indicates that the normal  
25 interaction between a chaperone and its natural ligand, i.e. the pilus subunit, is being affected either by being completely or substantially completely prevented, or by being inhibited, or expressed in another manner, reduced to a such an extent that the binding of pilus subunits to the chaperone  
30 is measurably lower than is the case when the chaperone is interacting with the pilus subunit at conditions which are substantially identical (with regard to pH, concentration of ions and other molecules) to the native conditions in the periplasmic space. Similarly, the enhancement of binding  
35 between the chaperone and the pilus subunit should be such that the binding of pilus subunits to the chaperone is measu-



rably higher than is the case when the chaperone is interacting with the pilus subunit at conditions which are substantially identical (with regard to pH, concentration of ions and other molecules) to the native conditions in the periplasmic space. Measurement of the degree of binding can be determined *in vitro* by methods known to the person skilled in the art (microcalorimetry, radioimmunoassays, enzyme based immuno assays, etc).

It should, on the basis of the above, be clear that prevention or inhibition of the normal interaction between a pilus subunit and a chaperone should have a substantially limiting effect on pilus assembly. However, an enhancement of the binding between pilus subunits and chaperones may also prove to be devastating to a bacterium. As will appear from example 2, different pilus subunits bind to *PapD* with different affinities and affecting this narrowly balanced system may also cause a limitation on the rate and efficiency of pilus assembly.

It is believed that even modest changes in the binding between pilus subunits and chaperones can have dramatic impact on the efficiency of pilus assembly, and thus on the ability of the bacteria to adhere. For example, if the change in the binding between a chaperone and one pilus subunit is such that the normal order of affinities between the chaperones and the pilus subunits which normally bind thereto is altered, then the normal assembly of the pilus should be disturbed, since the order of assembly of the pilus may be dependent *i.a.* on the affinities between the pilus subunits and the chaperone: The pilus subunits with the highest affinities to the chaperone may be incorporated before other pilus subunits with lesser affinities.

Thus, prevention, inhibition or enhancement of binding between pilus subunits and a periplasmic molecular chaperone have the effect of impairing pilus assembly, whereby the

infectivity of the microorganism normally expressing the pili is reduced.

Prevention, inhibition or enhancement of the binding between pilus subunits can be accomplished in a number of ways. A preferred method according to the invention of treatment and/or prophylaxis of diseases caused by tissue-adhering pilus-forming bacteria is to administer an effective amount of a substance to a subject in need thereof, the substance being capable of interacting with at least one type of molecular chaperone which binds pilus subunits during transport of these pilus subunits through the periplasmic space and/or during the process of assembly of the intact pilus, in such a manner that binding of pilus subunits to the molecular chaperone is prevented, inhibited or enhanced.

The substance can be any compound which has one of the above mentioned effects on the interaction between chaperones and pilus subunits and thereby on the assembly of the pilus. Especially interesting substances are those which are likely to interact with the pilus subunit binding part of the chaperone, but interaction with other sites in the chaperones may also cause prevention, inhibition or enhancement of the binding between pilus subunits and the chaperone. This can be an effect of direct steric blocking of the normal binding between the subunit and the chaperone, but it may also be an effect of a conformational change in the chaperone. A method of identifying substances to be used in the method of the invention is disclosed below.

The interaction between the substance and the chaperone may be a covalent as well as a non-covalent binding to the chaperone by the substance.

By the term "subject in need thereof" is in the present context meant a subject, which can be any animal, including a human being, who is infected with, or is likely to be

infected with, tissue-adhering pilus-forming bacteria which are believed to be pathogenic.

By the term "an effective amount" is meant an amount of the substance in question which will in a majority of patients have either the effect that the disease caused by the pathogenic bacteria is cured or, if the substance has been given prophylactically, the effect that the disease is prevented from manifesting itself. The term "an effective amount" also implies that the substance is given in an amount which only causes mild or no adverse effects in the subject to whom it has been administered, or that the adverse effects may be tolerated from a medical and pharmaceutical point of view in the light of the severity of the disease for which the substance has been given.

The route of administration of the substance could be any conventional route of administration, i.e. oral, intravenous, intramuscular, intradermal, subcutaneous etc. The oral route is preferred.

The dosage of such a substance is expected to be the dosage which is normally employed when administering antibacterial drugs to patients or animals, i.e. 1  $\mu\text{g}$  - 1000  $\mu\text{g}$  per kilogram of body weight per day. The dosage will depend partly on the route of administration of the substance. If the oral route is employed, the absorption of the substance will be an important factor. A low absorption will have the effect that in the gastro-intestinal tract higher concentrations, and thus higher dosages, will be necessary. Also, the dosage of such a substance when treating infections of the central nervous system (CNS) will be dependent on the permeability of the blood-brain barrier for the substance. As is well-known in the treatment of bacterial meningitis with penicillin, very high dosages are necessary in order to obtain effective concentrations in the CNS.

It will be understood that the appropriate dosage of the substance should be suitably be assessed by performing animal model tests, wherein the effective dose level (e.g. ED<sub>50</sub>) and the toxic dose level (e.g. TD<sub>50</sub>) as well as the lethal dose level (e.g. LD<sub>50</sub> or LD<sub>10</sub>) are established in suitable and acceptable animal models. Further, if a substance has proven efficient in such animal tests, controlled clinical trials should be performed. Needless to state that such clinical trials should be performed according to the standards of Good Clinical Practice.

Although the preferred way of preventing, inhibiting or enhancing the binding between pilus subunits and chaperones is to administer a substance with the above mentioned effects on the chaperone, other ways are possible. For instance, substances interacting with one type of pilus subunit could also have the effects described above, and for the same reasons. However, as the interaction with the chaperone is likely to exert effects on the assembly into the pilus of most, if not all, pilus subunits constituting the intact pilus, it is expected that the interaction with the chaperone will be the most efficient in terms of hampering bacterial infectivity.

As will appear from the examples below, most of the data on binding between chaperones and pilus subunits have been obtained by studying the interaction between the *PapD* chaperone from *E. coli*. However, since many tissue adhering bacteria have been found to express pili which share substantial homologies in their C-terminal part, and since substantial homologies have been demonstrated between the various periplasmic chaperones which until now have been isolated (see table A), it is justified to assume that some substances and classes of substances will be capable of interacting with the majority of existing periplasmic chaperones and thus be useful in the treatment and/or prophylaxis of diseases caused by the bacteria harbouring these chaperones when the sub-

stance is administered to patients infected with the bacteria.

Thus, the method of the invention for the treatment and/or prophylaxis is especially intended to be used in patients which are infected by bacteria selected from the group consisting of *Haemophilus* spp, *Helicobacter* spp, *Pseudomonas aeruginosa*, *Mycoplasma* spp, and all members of the Enterobacteriaceae family, including *Escherichia* spp, *Salmonella* spp, *Bordetella* spp, *Yersinia* spp, *Proteus* spp and *Klebsiella* spp. In this connection, especially the bacteria selected from the group consisting of *E. coli*, *Y. pestis*, *Y. enterocolitica*, *B. pertussis*, *K. pneumoniae*, *S. typhimurium*, *S. typhi*, *S. paratyphi*, *Helicobacter pylori*, *Proteus mirabilis* and *Haemophilus influenzae* are regarded as infectants which cause infections which can be treated and/prevented by the use of the method according to the invention.

Accordingly, in important aspects of the invention, the binding of a pilus subunit to a chaperone selected from the group consisting of *PapD*, *FimC*, *SfaE*, *FaeE*, *FanE*, *Cs3-1*, *Fl7D*, *ClpE*, *EcpD*, *Mrkb*, *FimB*, *SefB*, *HifB*, *MyfB*, *PsaB*, *PefD*, *YehC*, *MrpD*, *CssC*, *NfaE*, *AggD*, and *Caf1M* is prevented, inhibited or enhanced. It is especially preferred that the binding of *PapD* to at least one pilus subunit is affected.

As stated above, in a preferred embodiment of the invention, the prevention, inhibition or enhancement of the binding is accomplished by interacting with, in the molecular chaperone, a binding site which is normally involved in binding to pilus subunits during transport of these pilus subunits through the periplasmic space and/or during the process of pilus assembly.

As mentioned, in connection with the present invention, the binding motif between *PapD* and a peptide which constitute the 19 amino acids of the C-terminal of *PapG* (G1'-19'), a pilus subunit, has been determined. As is described in detail

herein, other chaperones share substantial homologies with *PapD* in this binding site. Thus, such a binding site is of great interest as a target for drugs which are intended to interact with periplasmic chaperones. Therefore, in a preferred embodiment of the above described methods of the invention, the binding site which is affected is one which binds G1'-19'.

Thus, an important aspect of the invention is a method as described above, wherein the binding site is a binding site to which the carboxyl terminal part of a pilus subunit is capable of binding, and which comprises site points substantially identical to the invariant residues Arg-8 and Lys-112 in *PapD*, and a polypeptide fragment which is capable of interacting with a  $\beta$ -strand of the carboxyl terminal part of the pilus subunit thereby stabilizing the binding of said subunit at the Arg-8 and Lys-112 site points of the binding site. An especially preferred aspect of the invention is a method as described above, where the binding site is the G-protein binding site of *PapD* as described herein.

The term "site-point" refers to a chemical group with well defined physical/chemical characteristics such as size, charge, hydrophobicity/hydrophilicity, polarity, direction of hydrogen bonds as well as a 3-dimensional position (distance and angle) relative to other such chemical groups. Thus, site-points which are "substantially identical" to Arg-8 and Lys-112, are chemical groups which substantially share the same well-defined physical/chemical characteristics as these two amino acids.

The term "invariant residues" refer to amino acid residues which can be found in a number of proteins without there being any variation with regard to the precise type of the amino acid and without there being any substantial variation in their function in the proteins. The presence of invariant residues in a large number of related proteins normally is an indication of the biological importance of such residues,

since mutations lacking these residues apparently lacks the function of the intact protein, too. As described herein, it has been found that all periplasmic molecular chaperones share the amino acid residues which are equivalent to Arg-8 and Lys-112. It is believed that these two residues therefore are of considerable importance to pilus-forming bacteria.

"A polypeptide fragment capable of interacting with a  $\beta$ -strand of the carboxyl terminal part of a pilus subunit" indicates that part of the chaperone (which is also a part of the binding site) is capable of interacting with a  $\beta$ -strand of the pilus subunit. This interaction serves as a stabilizing factor in the binding between the pilus subunit and the chaperone and is considered a very important part of the total motif of binding between the chaperone and the pilus subunit. Further, it has recently been rendered probable by the inventors that the  $\beta$ -strand serves as a template for the correct folding of the pilus subunit (cf. example 10).

As explained herein, the C-terminal part of many, if not all, known pilus subunits, share substantial homologies, which is another indication of the importance of the 3-dimensional structure of the pilus subunit as well as of the chaperone in order for the binding to take place and be stable.

As appears from the examples, another binding site residing in domain 2 of PapD has been identified. This binding site interacts with fusion protein MBP-G1'-140' as well as with a short peptide constituted of the C-terminal amino acid residues 125' to 140' of PapG. Thus, also this binding site is of great interest as a target for drugs which are intended to interact with periplasmic chaperones. Therefore, a preferred embodiment of the above described methods of the invention is a method wherein the binding site which is affected is one which binds either of the two above-described peptides.

It will be understood that the above-described methods comprising administration of substances in treating and/or

preventing diseases are dependent on the identification or de novo design of substances which are capable of exerting effects which will lead to prevention, inhibition or enhancement of the interaction between pilus subunits and periplasmic molecular chaperones. It is further important that these substances will have a high chance of being therapeutically active.

Thus, an aspect of the invention relates to a method for identifying a potentially therapeutically useful substance capable of interacting with a periplasmic molecular chaperone, thereby preventing, inhibiting or enhancing the interaction between a periplasmic molecular chaperone and a pilus subunit, the method comprising at least one of the following steps:

- 1) testing a candidate substance in an assay in which the possible prevention, inhibition or enhancement by the substance of the interaction between the periplasmic molecular chaperone and the pilus subunit is determined by
  - a) adding the substance to a system comprising the periplasmic molecular chaperone or an analogue thereof in an immobilized form and the pilus subunit or an equivalent thereof in a solubilized form and determining the change in binding between the pilus subunit or equivalent thereof and the periplasmic molecular chaperone or analogue thereof caused by the addition of the substance, or
  - b) adding the substance to a system comprising the pilus subunit or an equivalent thereof in an immobilized form and the periplasmic molecular chaperone or an analogue thereof in a solubilized form and determining the change in binding between the pilus subunit or equivalent thereof and the periplasmic molecular chaperone or analogue thereof caused by the addition of the substance, or



- 5 c) adding the substance to a system comprising the pilus subunit or an equivalent thereof as well as the periplasmic molecular chaperone or an analogue thereof in solubilized form and determining the change in binding between the pilus subunit or equivalent thereof and the periplasmic molecular chaperone or analogue thereof caused by the addition of the substance, or
- 10 d) adding the substance to a system comprising the pilus subunit or an equivalent thereof as well as the periplasmic molecular chaperone or an analogue thereof in solubilized form and measuring the change in binding energy caused by the addition of the substance, and identifying the substance as potentially therapeutically useful if a significant change in the binding energy
- 15 between the pilus subunit or equivalent thereof and the periplasmic molecular chaperone or analogue thereof is observed,
- 20 and identifying the substance as potentially therapeutically useful if a significant change in the binding or binding energy between the pilus subunit or equivalent thereof and the periplasmic molecular chaperone or analogue thereof is observed;
- 25 2) testing a candidate substance in an assay in which the possible prevention, inhibition or enhancement of the interaction between the periplasmic molecular chaperone and the pilus subunit is determined by
- 30 adding the substance to a system comprising living tissue-adhering pilus-forming bacteria followed by determination of the growth rate of the bacteria, a reduction in growth rate compared to a corresponding system wherein the substance has not been added being indicative of prevention, inhibition or enhancement of the binding between the periplasmic molecular chaperone and the pilus subunit, or

adding the substance to a system comprising living tissue-adhering pilus-forming bacteria followed by a determination of the tissue adhesion of the bacteria, a reduction in tissue adhesion compared to a corresponding system wherein the substance has not been added being indicative of prevention, inhibition or enhancement of the binding between the periplasmic molecular chaperone and the pilus subunit,

and identifying the substance as potentially therapeutically useful if a reduction in growth rate or tissue adhesion is observed after the addition of the substance; and

3) administering, to an experimental animal, a substance which has been established *in vitro* to prevent, inhibit or enhance the interaction between a periplasmic molecular chaperone and a pilus subunit, the experimental animal being inoculated with tissue-adhering pilus-forming bacteria before, simultaneously with or after the administration of the substance, and electing as a substance suitably capable of interacting with a periplasmic molecular chaperone, a substance preventing and/or curing and/or alleviate disease caused by the bacteria.

By the term "an equivalent of a pilus subunit" is meant a compound which has been established to bind to the chaperone in a manner which is comparable to the way the pilus subunit binds to the chaperone, e.g. by the demonstration of the pilus subunit and the equivalent competing for the binding to the chaperone. Preferred equivalents of pilus subunits are G1'-19'WT, MBP-G1'-140' and G125'-140', which are all described in detail herein.

The term "an analogue of a chaperone" denotes any substance which have the ability of binding at least one pilus subunit

in a manner which corresponds to the binding of said chaperone to a pilus subunit. Such an analogue of the chaperone can be a truncated form of the intact chaperone (e.g. one of the two domains of *PapD*) or it can be a modified form of the chaperone which may e.g. be coupled to a probe, marker or another moiety. Finally, the analogue of the of the chaperone can be an isolated, but partially or fully functional, binding site of the chaperone or a synthetic substance which mimics such a binding site.

- 10 The immobilisation mentioned above may be simple non-covalent binding to an adhering surface or a host or receptor molecule such as an antibody, or covalent binding to a spacer molecule such as a polymer or a peptide.

15 In the above mentioned step 1a), the pilus subunit or the equivalent thereof being bound to the periplasmic molecular chaperone or an analogue thereof can be detected in a number of ways, e.g. by the pilus subunit or the equivalent thereof being labelled, or by means of a labelled ligand (such as an antibody) capable of reacting with the pilus subunit or the  
20 equivalent thereof, or by means of a refractive index based determination of the extend of binding, such as the Pharmacia BiaCore® assay.

Accordingly, in step 1b) the periplasmic molecular chaperone or the analogue thereof being bound to the pilus subunit or  
25 the equivalent thereof may be detected by the periplasmic molecular chaperone or the analogue thereof being labelled, by means of a labelled ligand (e.g. antibody) capable of reacting with the periplasmic molecular chaperone or the analogue thereof, or by means of a refractive index based  
30 determination of the extend of binding, such as the Pharmacia BiaCore® assay.

In step 1c) the periplasmic molecular chaperone or the analogue thereof being bound to the pilus subunit or the equivalent thereof may be detected by separation of pilus sub-

unit/chaperon complexes (e.g. by ultracentrifugation, ultrafiltration, liquid chromatography, such as size exclusion chromatography, or electrophoresis). Described below is a method relying on the changes in fluorescence of a short  
5 PapG fragment when this fragment is bound to PapD. This method is a preferred assay in the method of the invention.

The determination of binding energy in step 1c) is preferably performed in a microcalorimetric system using the well-known technique of microcalorimetry.

10 The above-indicated steps serve 3 purposes. The types of assays in step 1) are intended to shed light over the ability of the candidate substance of interacting with the chaperone. In the instances wherein labelled substances, chaperones or antibodies are used, the label could be a radioactive label,  
15 a fluorescent or light absorbing label, an enzyme such as horse-radish peroxidase, a ligand such as biotin, or any other conventional labelling system known to the person skilled in the art. The detection of the labelled compound is then dependent on the choice of label: radioactivity may be  
20 measured in a liquid-scintillation counter, a gamma counter, or any other convenient detection system for radioactivity, enzyme-labels are detected by the presence or absence of a specific substrate for the enzyme (optical density assessment, chemical reactivity of the remaining substrate or of  
25 the product etc.), fluorescent labels may be detected by fluorescence microscopy or simple measurement of the fluorescent emission, light-absorbing labels may be detected by measurement of absorption of light of a characteristic wavelength, and biotin may be detected by its binding to strept-  
30 avidin.

The separation of high molecular complexes by ultracentrifugation or ultrafiltration in 1) may be detected by one of the components of the complex being labelled as described above; it is thus possible to detect the ratio between bound  
35 and unbound pilus subunit/equivalent, but the detection step

may also rely on the binding of antibodies to one of the components of the complex, and the subsequent detection of this antibody. Any conventional chromatographic technique may be employed (HPLC, FPLC, size exclusion, etc) The separation  
5 by electrophoresis may e.g. be performed by capillary electrophoresis.

The assays in step 2) all relate to the effects of the candidate substance on bacterial activity *in vitro*. The demonstration of a reduction in growth rate of the bacteria or a  
10 demonstration of reduced adherence to cells or synthetic surfaces in an assay of course cannot be contributed to the effect of interaction with chaperones only, but a demonstration of this kind should provide a good estimate of the potential therapeutical usefulness of such a substance.

- 15 The determination of growth rate may be performed by counting of colonies on solid agar plates striped with the bacteria, by counting bacterial density in liquid growth media (OD<sub>600</sub> determination), by measuring fluorescence of substances such as NAD(P)H, ATP, or amino acids, which are contained in the  
20 bacterial cells only, or by any other convenient detection system known to the person skilled in the art. The determination of adherence of the bacteria may be performed in a similar manner after the adhering bacteria have been isolated. A determination of adherence is preferably performed  
25 by measuring the ability of the bacteria to agglutinate red blood cells or receptor-coated latex beads, by measuring the bacterial adhesion to receptor-coated microtiter plates, or by measuring the bacterial adhesion to other synthetic surfaces.
- 30 In a preferred embodiment of the method described above, the living, tissue adhering pilus-forming bacteria are of a protease deficient strain, the protease being one which is at least partially responsible for the degradation of pilus subunits. One especially preferred type of strain is the  
35 degP41 strain of *E. coli*. As described herein, the degP41

strain lacks activity of the DegP protease which is responsible for degradation of pilus subunits in the *E. coli* when these are not rescued into periplasmic space by *FapD*, and *degP41* strains are thus especially sensitive to changes in the efficiency of *PapD*, as the accumulation of pilus subunits is toxic to the cell. It is believed that equivalent proteases exist in other pilus expressing bacteria.

The animal study in step 3) is performed in order to demonstrate the potential therapeutic usefulness of the candidate substance *in vivo*. Further, as already mentioned above, such animal studies should also establish the *a priori* values regarding effective dosage and toxicity before the candidate substance finally is tested in human beings in controlled clinical trials. The animal studies should also provide information regarding the convenient formulation of the substance in a pharmaceutical preparation as well as the preferred route of administration, as it is possible to obtain, from the animal model, data for absorption of the substance as well as data for the metabolism and excretion of the substance. The experimental animal is preferably a mouse, a rat, a cat, a dog, a monkey, a horse, a cow, a pig, or a chicken.

The term "suitably capable of interacting with a molecular chaperone" is intended to indicate that a substance, apart from being capable of interacting with a molecular chaperone, also is capable of exerting effects in an *in vivo* system, i.e. that the substance in addition to its binding capability also exhibits compatibility with a biological system, i.e. a patient.

Although the above-indicated *in vivo* studies, especially the experiments in animal models, are the best indicators of the potential therapeutical usefulness of a substance in the prevention, inhibition, or enhancement of the binding between a chaperone and a pilus subunit, it should not be forgotten that the *in vitro* assays outlined above serve as important

leads when developing compounds with a therapeutical potential. If one relied only on *in vivo* assays, it is very likely that compounds which in fact exhibit the desired effect on the chaperone/pilus subunit interaction would be screened out  
5 by the *in vivo* assays, because these compounds could lack e.g. the ability to penetrate biological membranes. When using the *in vitro* assays, a much greater chance of finding a lead compound is maintained.

The evaluation of the effect of a substance tested in the *in vitro* assays described herein (cf. in this connection especially the examples) depends on a number of factors. It will be understood by the skilled person that a small molecule could be added in rather high molar concentrations in order to exert an effect on the chaperone/pilus subunit interaction  
10 (and even then the small molecule may still be an interesting lead compound), whereas larger molecules may exert marked effects even in rather low molar concentrations. In general, when any *in vitro* assay described herein is regarded as having a positive result when testing a candidate substance  
15 (i.e. that the substance tested shows a "significant" effect), the following condition should be fulfilled: The compound should exert a significant effect on pilus subunit/chaperone interaction (or on an interaction in an equivalent system which correlates well to pilus subunit/chaperone interaction), the significant effect being one  
20 which with no doubt can be attributed to the interaction between the substance and the chaperone and which is not an unspecific interaction between the chaperone and the substance (due to e.g. radical changes in the physical and  
25 chemical environment when the substance is added). One way of excluding unspecific interactions as the reason for the exerted effect is to use at least one control which is a chemically comparable substance (with respect to molecular mass, charge/polarity and gross 3-dimensional conformation  
30 (globular, fibrillar etc.)). If the control does not result in substantially the same effect in the assay as the substance,  
35

it can be concluded that the substance must be regarded as an assay positive substance.

The assays described in the examples are all good examples of assay types, which could serve as the test system in the above-described method of the invention. However, it is preferred that the method described in example 10 employing a fluorescence labelled variant of a pilus subunit is used in step 1c). This assay may shortly be described as follows:

- 10 - adding the substance to a first system comprising the periplasmic molecular chaperone or an analogue thereof,
- subsequently adding a pilus subunit or an equivalent thereof which has been labelled with an environmentally sensitive fluorescent probe,
- 15 - determining the fluorescent emission at a particular wavelength which is indicative of the amount of binding between the periplasmic molecular chaperone or the analogue thereof and the pilus subunit or the equivalent thereof, and
- comparing the determined fluorescent emission to fluorescent emission determined in a corresponding second system containing substantially the same concentrations of the molecular chaperone or the analogue thereof and the pilus subunit or the equivalent thereof but substantially no substance,

25 a significant difference in fluorescent emission between the first and second system being indicative of interaction between the periplasmic molecular chaperone or the analogue thereof and the substance.

The advantage of this assay is that it may be employed for quantitative determinations of the effect of the tested substance on the chaperone/pilus subunit system. By using this assay the inventors have e.g. determined the constant of



binding between a *PapG* analogue and *PapD*. The quantitative determinations may be performed by performing the determination of fluorescent emission in the second system a plurality of times at varying molar ratios between the pilus subunit or the equivalent thereof and the periplasmic chaperone and the equivalent thereof, whereupon the constant of binding between the pilus subunit or equivalent thereof and the periplasmic molecular chaperone or analogue thereof is assessed from the determined fluorescent emission data. From the data obtained in this way it is also possible to determine the binding constant of the substance in a parallel manner, which will appear from claim 11.

It will be understood that the above-indicated method for identifying a potentially therapeutically useful substance is dependent on the actual presence of the substance. Normally, it is necessary to either purify or synthesize the candidate substance before it is subjected to the above-mentioned method. However, since many such candidate substances are likely to be tested before a substance which is suitably capable of interacting with a chaperone will be identified, it is of interest to identify such substances before they are subjected to the method above, thereby diminishing the resources spent on purification and/or synthesis steps.

Hence, the invention also relates to a method for identifying and/or designing a substance, X, capable of interacting with a chaperone, e.g. binding to the chaperone, with a predicted binding energy equal to or better than a predetermined threshold value, the method comprising

- 1) selecting a substance, A, which could potentially interact with a site in the chaperone, and providing a 3-dimensional structural representation thereof,
- 2) predicting the binding free energy between the substance A and the site in the chaperone,

- 3) if the predicted binding free energy between the substance A and the site in the chaperone is equal to or better than the predetermined threshold value, then identifying the substance A as the substance X,
- 5 4) if the predicted binding free energy between the substance A and the site in the chaperone is not equal to or better than a predetermined threshold value, then modifying the 3-dimensional structural representation and predicting the binding free energy between the thus  
10 modified substance, B, and the site in the chaperone, and
- 5) repeating step 4 until the predicted binding free energy determined between the resulting substance, X, and the site in the chaperone is equal to or better than the predetermined threshold value.
- 15 It is possible to expand the above-mentioned method with two further steps, wherein the actual binding free energy is determined, in order to establish that the experimental binding free energy also is better than the predetermined threshold value. By performing the following two steps
- 20 6) providing a sample of the chemical substance X and a sample of the chaperone and measuring the binding free energy between the chemical substance X and the chaperone (e.g. by microcalometry as mentioned above), and establishing that the measured binding free energy between the  
25 chemical substance X and the chaperone is equal to or better than the predetermined threshold value, and optionally
- 7) subjecting the substance X to the method mentioned above for identifying a substance suitably capable of interacting with a chaperone, in order to verify that the substance X is a potentially therapeutically useful substance capable of interacting with a chaperone,  
30

it is thus verified that the binding free energy between the candidate substance and the chaperone actually is better than the predetermined threshold value. Step 7) further establishes that the candidate substance stands good chances of being therapeutically useful.

The phrase "predicting the binding free energy" is meant to imply that the binding free energy is determined by calculation rather than by performing experimental work determining the actual binding free energy. One (theoretical) way of predicting binding free energy is by performing free energy perturbation (FEP) calculations on the interacting substances, but because of the vast amount of calculations such an approach would have as a result it is preferred that the empirical approximative method described below is employed.

The term "better than" is intended to mean that the binding free energy has a value which is higher than the binding free energy which has been chosen as the threshold value, meaning that the  $\Delta G$  is numerically higher than the threshold value selected. Or in other words: The term is intended to mean that the binding between the substance and the chaperone is more favourable energetically than the situation were the substance and the chaperone are suspended independently in solution.

In order to predict the binding energy in the above-indicated method, according to the invention it is especially preferred to use the following method:

Assessing the average energy difference,  $\langle \Delta V_{X-s}^{el} \rangle$ , defined as  $\langle V_{X-s}^{el} \rangle_B - \langle V_{X-s}^{el} \rangle_A$ , between the contribution from polar interactions to the potential energy between the chemical substance X and its surroundings (denoted s) in two states, one state (A) being where the chemical substance is surrounded by solvent, the other state (B) being where the chemical substance, bound to a periplasmic molecular chaperone or an analogue thereof, is surrounded by solvent,

assessing the average energy difference,  $\langle \Delta V_{X-s}^{vdw} \rangle$ , defined as  $\langle V_{X-s}^{vdw} \rangle_B - \langle V_{X-s}^{vdw} \rangle_A$ , between the contribution from non-polar interactions to the potential energy between the chemical substance X and its surroundings (denoted s) in two states, one state (A) being where the chemical substance is surrounded by solvent, the other state (B) being where the chemical substance, bound to a periplasmic molecular chaperone or an analogue thereof, is surrounded by solvent, and

calculating the absolute binding free energy as an adjusted combination of the two above-mentioned average energy differences.

In the mathematical equations herein, the symbol  $\langle \rangle$  means molecular dynamics average. The index X-s means compound-solvent (or compound-surrounding), the letter "X" denoting the chemical substance X. Normally the substance X will function as an inhibitor of the binding between the periplasmic chaperone and pilus subunits, but as discussed herein, it is also a possibility that the compound or drug will affect the chaperone in such a way that the binding between pilus subunits and the chaperone is enhanced. The superscript "el" designates the polar or electrostatic energy, while the superscript "vdw" indicates "van der Waals", another designation for the non-polar interactions. The symbol  $\Delta$  indicates that the quantity in state A is subtracted from the quantity in state B.

In the present context the term "an analogue of a periplasmic molecular chaperone" should be understood, in a broad sense, any substance which mimics (with respect to binding characteristics) an interesting part of a periplasmic molecular chaperone (e.g. the pilus subunit binding part(s)), and the interaction of which with a chemical substance or a group or plurality of chemical substances, e.g. drug candidates, is to be studied. Thus, the analogue may simply be any other chemical compound regarded as capable of interacting with the chemical substance in a manner which mimics the binding

between the chaperone and a pilus subunit in vivo, but most often the analogue will be a relatively large molecule, in other words a macromolecule such as a protein or an oligonucleotide, which is relatively large compared to the chemical substance; although the chemical substance interacting with the analogue, of course, in itself be a macromolecule. In the present context, the periplasmic molecular chaperone or analogue thereof is preferably the periplasmic chaperone or an analogue thereof which exhibits at least one interesting binding characteristic relevant for the assembly of pili.

The basis for the above-indicated approach for determining the binding free energy is explained in the following:

As a starting point is taken the linear response approximation for electrostatic forces which for polar solutions as a result yields quadratic free energy functions in response to the development of charges. This is, e.g., the familiar result from Marcus' theory of electron transfer reactions (Marcus, 1964). For a system with two states, A and B, given by two potential energy functions  $V_A$  and  $V_B$  one obtains, within the approximation of harmonic free energy functions of equal curvature, the relationship (see Lee et al., 1992 and references therein):

$$\lambda = \langle V_B - V_A \rangle_A - \Delta G_{AB} = \langle V_A - V_B \rangle_B + \Delta G_{AB} \quad (a)$$

where  $\Delta G_{AB}$  is the free energy difference between B and A,  $\lambda$  the corresponding reorganisation energy and  $\langle \rangle_i$  denotes an average evaluated near the minimum of the potential  $i$ . Thus,

$$\Delta G_{AB} = \frac{1}{2}(\langle \Delta V \rangle_A + \langle \Delta V \rangle_B) \quad (b)$$

where  $\Delta V$  now denotes the energy difference  $V_B - V_A$ . If the hydration of a single ion is considered, this can be shown to give  $\Delta G_{sol}^{el} = \frac{1}{2} \langle V_{X-S}^{el} \rangle$ , i.e. that the electrostatic contribution to the solvation energy equals half of the corresponding ion-solvent interaction energy (Warshel and Russell, 1984; Roux

et al., 1990). Returning now to the binding problem, this result may be exploited in the following manner: For each solvation process, i.e. solvation of the substance in water and inside the protein, two states are considered where the first has the substance in vacuum and a non-polar cavity (given, e.g., by Lennard-Jones potential) already made in the given environment. The second state corresponds to the intact substance surrounded by water or the solvated protein. The linear response approximation will then again give that

10  $\Delta G_{bind}^{el} \approx \frac{1}{2} \langle V_{X-S}^{el} \rangle$ , where  $V_{X-S}^{el}$  is the solute-solvent electrostatic term. Hence, the electrostatic contribution to the binding free energy can be approximated by  $\Delta G_{bind}^{el} \approx \frac{1}{2} \langle V_{X-S}^{el} \rangle$  (where the  $\Delta$  now refers to the difference between protein and water) and thus obtained from two MD simulations of the

15 solvated substance and of the substance-protein complex.

The validity of the linear response results in the case of ionic solvation has been confirmed, e.g., in the study by Roux et al. (1990). Some additional calculations were also performed on simple systems that corroborate the approximation of equation b. These tests were carried out by comparing

20 the free energy obtained from FEP/MD simulations of charging  $Na^+$  and  $Ca^{2+}$  ions in a spherical water system (Åqvist, 1990) with the corresponding  $\langle V_{X-S}^{el} \rangle$  from 75 ps MD trajectories. This yielded factors relating  $\langle V_{X-S}^{el} \rangle$  to  $\Delta G_{sol}^{el}$  of 0.49 for  $Na^+$

25 and 0.52 for  $Ca^{2+}$ , both values being close to the predicted result of  $\frac{1}{2}$ . A similar test on the charging of a methanol molecule, given by the OPLS potential (Jorgensen, 1986) in water gave a  $\Delta G_{sol}^{el} / \langle V_{X-S}^{el} \rangle$  ratio of 0.43.

A crucial question is how to account for the contribution of

30 non-polar interactions and hydrophobic effects to the free energy of binding which was termed  $\Delta G_{bind}^{vdw}$ . In the ideal case, it should be possible to estimate this contribution from the non-polar (or van der Waals) interaction energies. The liquid theories of Chandler and coworkers (Chandler et al., 1983;

35 Pratt and Chandler, 1977) have been successfully used to analyze hydrophobic effects and to calculate free energies of

transfer for some non-polar molecules (Pratt and Chandler, 1977), but no analytical treatment of that kind seems possible for solvation in an inhomogeneous environment such as a protein's active site. However, it has been noted that the experimental free energy of solvation for various hydrocarbon compounds, such as n-alkanes, depends approximately linearly on the length of the carbon chain both in their own liquids as well as in water (Ben-Naim and Marcus, 1984). MD simulations of n-alkanes solvated in water and in a non-polar van der Waals solvent have been carried out, which indicate that also the average solute-solvent interaction energies vary approximately linearly with the number of carbons in the chain (the relationships being different in different solvents, of course). It thus seem possible that a simple linear approximation of  $\Delta G_{bind}^{vdw}$  from  $\langle \Delta V_{X-S}^{vdw} \rangle$  might be able to account for the non-polar binding contribution. For instance, if  $\sigma$  is considered some appropriate measure of the size of the solute and if the solute-solvent van der Waals interaction energies and the corresponding non-polar free energy contributions (both in water and protein) depend linearly on  $\sigma$ , such that

$$\langle V_p^{vdw} \rangle = \alpha_p \sigma, \quad \langle V_w^{vdw} \rangle = \alpha_w \sigma, \quad \Delta G_p^{vdw} = \beta_p \sigma \text{ and } \Delta G_w^{vdw} = \beta_w \sigma$$

then  $\Delta G_{bind}^{vdw} = \frac{\beta_p - \beta_w}{\alpha_p - \alpha_w} \langle \Delta V_{X-S}^{vdw} \rangle$  is obtained. Since it seems difficult to derive a factor relating the two quantities in a reliable way from purely theoretical considerations, the approach is taken to empirically try to determine such a relationship which is capable of reproducing experimental binding data. Thus, the free energy of binding is in one embodiment of the invention approximated by

$$\Delta G_{bind} = \frac{1}{2} \langle \Delta V_{X-S}^{el} \rangle + \alpha \langle \Delta V_{X-S}^{vdw} \rangle \quad (1)$$

the parameter  $\alpha$  being determined by empirical calibration.

Although, as discussed above, a theoretical prediction of the coefficient for  $\langle \Delta V_{X-S}^{el} \rangle$  is  $\frac{1}{2}$ , it may be practically useful to

also treat this coefficient as an empirical parameter. This would lead to the free energy of binding being approximated by

$$\Delta G_{bind} = \beta \langle \Delta V_{X-S}^{el} \rangle + \alpha \langle \Delta V_{X-S}^{vdw} \rangle \quad (1b)$$

- 5 where both parameters,  $\alpha$  and  $\beta$ , are determined by empirical calibration.

Finally, in some cases, it seems suitable to add an additional constant term to Equation 1, so that the equation becomes

$$10 \quad \Delta G_{bind} = \beta \langle \Delta V_{X-S}^{el} \rangle + \alpha \langle \Delta V_{X-S}^{vdw} \rangle + c \quad (2)$$

- where  $c$  is a constant reflecting extrapolation to zero size of the chemical substance, that is, where the regression line is distinctly offset from origin when moving towards zero size of the chemical substance. The parameter  $c$  may also be  
 15 used to correct for possible systematic errors due to e.g. the neglect of induced polarisation, possible force field deficiencies etc. In these cases,  $c$  will normally assume a value between -10 and 10 kcal/mol, typically between -3 and 3 kcal/mol, such as between -2 and 2 kcal/mol, e.g. between -1  
 20 and 1 kcal/mol. However, it is anticipated that in many cases,  $c$  can suitably be set to zero, as the extent of deviation will be of minor importance for the usefulness of the predicted values.

- If also the electrostatic coefficient is treated as an empirical  
 25 cal parameter, the approximation of the binding free energy assumes its most general form, namely

$$\Delta G_{bind} = \beta \langle \Delta V_{X-S}^{el} \rangle + \alpha \langle \Delta V_{X-S}^{vdw} \rangle + c \quad (2b)$$

where now both  $\alpha$ ,  $\beta$  and  $c$  are to be determined by empirical calibration.



While the solvent used in the above method is suitably and most often an aqueous solvent like water, it is within the scope of the invention to take any other suitable solvent as a starting point, including, e.g., methanol, ethanol, acetone, acetonitrile, chloroform, hexane, etc., or mixtures thereof or combinations of such solvents or mixtures thereof with water. The selection of the solvent will be of little importance to the predicted values as long as the solvent is one which is able to dissolve or solvate the receptor molecule and the substance (in the present context this means that a sufficient amount of the periplasmic molecular chaperone or analogue thereof can be homogeneously mixed with the solvent without precipitation so as to allow the determination of binding energies by some suitable method), but there may be cases where it is advantageous to modify the solvent environment (e.g. by modulating the ionic strength) in which the interaction of the substance and the receptor molecule is to take place. If the environment in which the interaction between the chemical substance, such as a drug, and a periplasmic molecular chaperone or an analogue thereof is to take place in the actual use of the drug is the human body, it might be particularly suitable to imitate e.g. human plasma as the solvent.

A thorough discussion of the above-referenced method for determining the binding free energy between two molecules can be found in International Patent Application No. PCT/IB94/00257 and in Åqvist et al, 1994. These two documents are hereby incorporated by reference.

The above referenced method for determining the binding free energy has been employed in example 3 in order to identify compounds which stands a high chance of binding to the binding site of PapD; this means that calculations as the above described have been performed as the last theoretical step before compounds have actually been synthesized.

It will be understood that the above mentioned methods for identifying substances capable of interacting with chaperones will prove especially efficient in identifying substances which are of potential pharmaceutical value if the site to which they bind is known to be involved in pilus assembly.

Therefore, it is preferred that the site with which the substance may potentially interact, and to which the binding free energy is predicted, is the pilus subunit binding part of a molecular chaperone, such as the pilus subunit binding site of a molecular chaperone selected from the group consisting of *PapD*, *FimC*, *SfaE*, *FaeE*, *FanE*, *Cs3-27*, *F17D*, *ClpE*, *EcpD*, *Mrkb*, *FimB*, *SefB*, *HifB*, *MyfB*, *PsaB*, *PefD*, *YehC*, *MrpD*, *CssC*, *NfaE*, *AggD* and *Caf1M*, or an analogue of such a pilus subunit binding site, since the pilus subunit binding sites in these chaperones show extensive homologies. It is especially preferred that the binding site is the pilus subunit binding site of *PapD* or an analogue thereof.

As will appear from the examples, an important part of the chaperone binding motif has been discovered and a peptide corresponding to this motif has been synthesized and co-crystallized with *PapD* to provide a structural basis for the mechanism of action of *PapD*. The molecular details of the *PapD*-adhesin recognition interface clearly demonstrate the function of the conserved cleft in the entire pilus chaperone superfamily in subunit binding and in shuttling virulence determinants to the surface of pathogenic bacteria. The *PapD*-peptide crystal structure essentially represents a "snapshot" of a fundamental process in bacterial pathogenesis: the interaction of an adhesin with a chaperone, which is a prerequisite to adhesin presentation on the microbial surface.

Thus, the inventors of the present invention have by the use of X-ray crystallography elucidated the mechanism of binding between *PapD* and the pilus subunit *PapG* thereby identifying an essential part of a defined binding site responsible for the binding between pilus subunits and their periplasmatic

chaperones, and thus providing a method to enable drug design of chaperone inhibiting anti-bacterial compounds.

Having determined the location of a promising binding site for inhibitory ligands as described above (see details in examples 1 and 2), the computer programs "PLIM" and "PLIM\_DBS" (developed by Symbicom AB) have been used to find templates for families of compounds capable of binding to the binding site.

PLIM is a Protein Ligand Interaction Modeller that constructs putative ligands for a protein using thermodynamic criteria. It calculates the energy of interaction between the protein and sample probes that are successively placed at different points on a regular grid around the molecule. For each position and orientation the interaction energy between the probe and the atoms of the protein is calculated. The energies are stored, and the best positions for a particular probe are written out (the basic calculations are described by Goodford (1985) and Boobbyer (1989) and implemented in the commercially available program GRID; the PLIM implementation is somewhat different in that the energy values are converted to discrete points that are associated with the chemical probe, enabling easy output to e.g. data base searching programs). The program then builds up the ligand by incorporating selected probe atoms at positions of energy minima on the grid. The user selects which atoms and groups should be used as probes, and which criteria should be used to determine those that will be incorporated into the ligand. The energy is calculated as the sum of electrostatic, Van der Waals and hydrogen-bonding contributions as described herein.

The PLIM runs result in a number of suggested positions and orientations of favourable chemical groups in the region near the binding site. These groups which have physical properties like charge, hydrogen bonding directionality and extended atom radia, will hereafter be denoted "site points".

A search for potential ligands is then made by searching a database for known molecular structures that match the positions of these groups of site points, using PLIM\_DBS.

The core of PLIM\_DBS is an algorithm for subgraph isomorphism (cf. Ullman (1976) and Brint (1987)), where three sitepoints are represented as a distance matrix ("the pattern matrix"). The program looks for this distance pattern in the distance matrix formed from every entry in the database. If the pattern is found, the entry is superimposed on to the sitepoints and if the corresponding atom types match the entry and its orientation is saved in a hit-list. Added to this basic scheme a number of options regarding surface complementarity can be used, i.e. only entries which are matching the protein surface with respect to hydrophobic and steric properties are saved.

PLIM\_DBS is thus a database searcher which hunts through a collection of 3-dimensional molecule coordinate sets, looking for entries that contain a certain pattern of atoms. This pattern is specified in terms of atom type, and of spatial position and orientation; for instance a search may be made for compounds containing an sp<sup>3</sup> carbon atom that is 4.2 Å from a sp<sup>2</sup> oxygen and 5.1 Å from a hydroxyl group that in turn is 5.6 Å from the oxygen. The strictness of the search can be adjusted by the user by varying the tolerance on the distance criteria and the atom-type matching, determining, for instance, whether a sp<sup>2</sup> carbon that is a little more than 4.2 Å from an oxygen should be considered as a hit. Those hits that are found are then ranked according to a score that reflects how well the target atoms superimpose on the real molecule, and also on how complementary the molecular surface of the compound is to that of the binding pocket of the protein.

The result from a PLIM\_DBS search is a list of molecular structures and their atomic coordinates, superpositioned on to the sitepoints, and given a score ("goodness of fit").

The procedure does not try to optimize the positioning of the structures, nor does it perform any molecular mechanics or dynamics calculations. Both protein and the extracted structures are treated as rigid bodies.

- 5 The structures from the database search are displayed in the context of the protein and its surface on a graphics system using a commonly available molecular modelling package. Usually the structures show some unfavourable interactions with the protein, or lack groups to fill out e.g. hydrophobic  
10 pockets. Hence, the structures from the database search are regarded as templates, to be modified and improved by an organic chemist. This process also involves choosing compounds which are easy to synthesize, which is of particular interest if the synthesis capacity is limited.
- 15 The best of these database hits are thus examined visually using a computer-graphic modelling system, and the most promising of these are selected according to a wealth of physico-chemical reasoning.

- The templates are modified using a small molecule 3D builder  
20 (MacMimic). Each template gives rise to a compound class, e.g. denoted "hdo". Each modification assigned a specific number (e.g. hdo\_3) and the coordinates and a description are stored in a tree structure, using the program ARVS\_JAKT developed by Symbicom. The design is performed in a collabo-  
25 ration between protein structure experts and organic chemists, in order to provide the best tools possible for the chemists who will actually synthesize the compounds.

- The efficacy of these modifications is finally assessed using molecular-dynamics free energy calculations as described  
30 herein to study the stability of the protein-ligand complex (Åqvist et al., 1994; Åqvist and Medina, 1993).

In order to maximize the efficiency of the above-mentioned methods for identifying/designing substances which are

capable of interacting with a molecular chaperone, it is preferred that the substance A is likely to be a substance which is capable of binding to the selected binding site.

In view of the above-described *modus operandi* for selecting  
5 substances which should interact with chaperones like *PapD*, this can, according to the invention, be accomplished when the substance A is selected by performing the following steps:

- 10 - co-crystallizing the periplasmic molecular chaperone or the analogue thereof with a ligand capable of interacting with a site in the periplasmic molecular chaperone or the analogue thereof and establishing the three-dimensional conformation of the periplasmic molecular chaperone or the analogue thereof and the ligand when interacting by  
15 means of X-ray crystallography,
- using the above-established conformation of the periplasmic molecular chaperone or the analogue thereof to establish a 3-dimensional representation of the site in the periplasmic molecular chaperone or the analogue thereof  
20 interacting with the ligand during binding,
- selecting a number of distinct chemical groups, X1, and determining the possible spatial distributions of the X1 chemical groups which maximizes the binding free energy between the chemical groups and the site in the chaperone  
25 or the analogue interacting with the ligand,
- extracting, from a database comprising three-dimensional representations of molecules, a molecule which has the X1 chemical groups in the possible spatial distributions determined above,
- 30 - optionally modifying the 3-dimensional representation of the molecule extracted from the database, and

identifying the optionally modified molecule as the substance A.

According to the invention the above-indicated steps are especially preferred when the ligand is a pilus subunit or a part thereof with which the chaperone normally interacts during transport of the pilus subunit through the periplasmic space and/or during pilus assembly.

By the use of the above-mentioned method for identifying substances capable of interacting with the periplasmic chaperones, several classes of substances have been identified which have proved promising in the design phase.

Drug design efforts for inhibitors/enhancers to *PapD* have concentrated on the region of the molecule where the G-peptide is observed to bind. This region will now be described in detail, using one of the putative inhibitors (bpy\_9, see below) as a reference structure.

The binding site is dominated by the central charged side chains of Arg-8 and Lys-112, which bind to the sulphate moiety of bpy\_9. Adjacent to this is a small, shallow hydrophobic pocket formed by the side-chains of Ile-154, Ile-194, and Thr-7, against which the 2-ethyl group of bpy\_9 packs (Pocket 1). A group as large as a phenyl ring could be accommodated here, attached to sugar position 2, and with possible substituents that could receive or donate a hydrogen bond to Thr-7, or donate to the backbone carbonyl of 198.

There is also a larger pocket comprising residues Leu-4, Ile-111, Thr-7 and Thr-109, in which the phenyl ring of bpy\_9 nestles (Pocket 2). A group as large as naphthalene could be substituted on the 6-position of the carbohydrate scaffold to fill this sub-site, with substituents that could form a hydrogen bond to Thr-109, or to any of the polar backbone atoms of residues Leu-4, Arg-6, or Lys-110. Then there is a long, shallow patch that includes Tyr-87, and the aliphatic

regions of Lys-110 and Lys-112, which could accommodate a tricyclic system such as 2-phenanthryl substituted on position 3 of the sugar (Patch 3). Substituents to hydrogen bond to Tyr-87 or the backbone of Lys-110 can be considered, as well as negatively charged groups to complement the side chains of Lys-110 and Lys-112. Tyr-87 is a potential charge-transfer donor to a electron deficient  $\pi$ -system such as a nitroaryl.

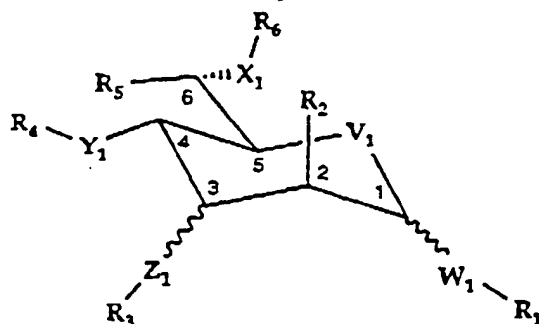
Models of several homologous chaperones (*SfaE*, *MrkB*, *HifB* and *FimC*) have been made from the *PapD* structure, and differences between the model structures has influenced the design of inhibitors, such that proposed ligands should bind to all the structures looked at. For example, it is tempting to complement the charged side chain of Arg-200 with an acidic group in the ligand, but since two of the other structures have an Asp at this position, the residue is not considered as a good candidate. Arg-8 is fully conserved, as is Lys-112, Thr-7 and Ile-11. Tyr-87 becomes a Trp in three structures, but the overall nature of patch 3 is not changed by this. *PapD* is actually alone with its Thr-Lys sequence at 109-110, all of the four other structures being Ser-Arg here, but again, these conservative changes do not significantly alter design criteria.

Although substances which interact with the binding site responsible for binding to G-proteins are obvious candidates as inhibitors/enhancers of periplasmic chaperones, it will be understood that molecules capable of interacting with other sites in periplasmic chaperones are interesting in this aspect, too. It is highly possible that an interaction with another site than the one binding G-protein in i.e. *PapD* may cause *PapD* to either be prevented, inhibited or enhanced in its action as a periplasmic chaperone.

As mentioned above, a family of substances (called the bpy family herein) is an important aspect of the invention. Thus,



the invention relates to novel compounds of the general formula:



I

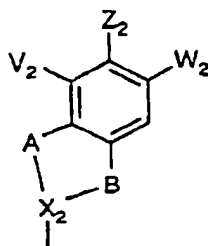
wherein

$V_1$  is O, S, SO, SO<sub>2</sub>, CH<sub>2</sub>, C(OH)H, CO or CS;

5  $W_1$  is O, S, SO<sub>2</sub>, SO<sub>3</sub>, CH<sub>2</sub> or NH;

$R_1$  is H; C<sub>1-24</sub> alkyl, C<sub>1-24</sub> alkenyl or C<sub>1-24</sub> alkynyl, which alkyl, alkenyl and alkynyl may be substituted with one or more substituents independently selected from OH, -CONH<sub>2</sub>, -CSNH<sub>2</sub>, -CONHOH, -CSNHOH, -NHCHO, -NHCONH<sub>2</sub>, -NHCSNH<sub>2</sub>,  
 10 -NHSO<sub>2</sub>NH<sub>2</sub> and -SO<sub>2</sub>NH<sub>2</sub>; acyl; or -(CH<sub>2</sub>CH<sub>2</sub>O)<sub>s</sub>-H, wherein s = 1, 2, 3;

$R_2$  is a group of the formula



wherein

A is -CH-(CH<sub>2</sub>)<sub>n</sub>-, or -CH=CH-(CH<sub>2</sub>)<sub>n-1</sub>- (n>0) or

$$-\text{CH}-(\text{CH}_2)_n-;$$

$$\parallel$$

$$\text{Y}_2$$

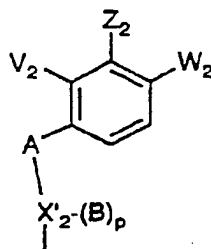
B is  $-(\text{CH}_2)_m-$  or  $=\text{CH}-(\text{CH}_2)_{m-1}-$  ( $m>0$ );

5  $\text{X}_2$  is N, CH or C (when B is  $=\text{CH}-(\text{CH}_2)_{m-1}-$ ; and

$\text{Y}_2$  is O, S, NH,  $\text{H}_2$  or H ( $n=1$ ); and

$4>m+n>0$ ,  $n<3$ , and  $m<3$ ;

or  $\text{R}_2$  is a group of the formula



wherein

10 A is  $-\text{CH}-(\text{CH}_2)_n-$ , or  $-\text{CH}=\text{CH}-(\text{CH}_2)_{n-1}-$  ( $n>0$ ) or  $-\text{CH}-(\text{CH}_2)_n-$ ;

$$\parallel$$

$$\text{Y}_2$$

B is  $-(\text{CH}_2)_m-$  or  $=\text{CH}-(\text{CH}_2)_{m-1}-$  ( $m>0$ ); and

15  $\text{X}'_2$  is O, NH,  $\text{CH}_2$  or S (when  $p=0$ ); N or CH ( $p=1$ ); or C (when  $p=1$  and B is  $-\text{CH}-(\text{CH}_2)_{m-1}-$ );

20  $\text{V}_2$ ,  $\text{Z}_2$  and  $\text{W}_2$  are independently H, OH,  $-\text{CONH}_2$ ,  $-\text{CSNH}_2$ ,  $-\text{CONHOH}$ ,  $-\text{CSNHOH}$ ,  $-\text{NHCHO}$ ,  $-\text{NHCONH}_2$ ,  $-\text{NHCSNH}_2$ ,  $-\text{NHSO}_2\text{NH}_2$ ,  $-\text{SO}_2\text{NH}_2$ , or  $\text{V}_2$  and  $\text{Z}_2$ , or  $\text{Z}_2$  and  $\text{W}_2$  together form  $-\text{NHC}(\text{O})\text{NH}-$ ,  $-\text{C}(\text{O})\text{NHC}(\text{O})-$ ,  $-\text{NHS}(\text{O}_2)\text{NH}-$ ,  $-\text{C}(\text{O})\text{NHO}-$ ,  $-\text{C}(\text{S})\text{NHO}-$ ,  $-\text{S}(\text{O}_2)\text{NHO}-$ , or  $-\text{S}(\text{O}_2)\text{NHC}(\text{O})-$ ;

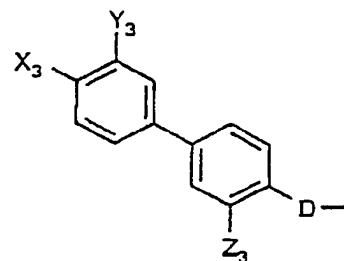
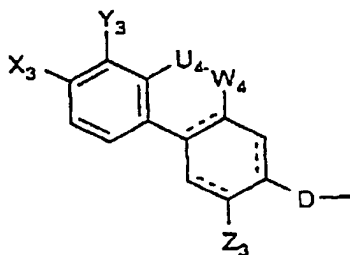
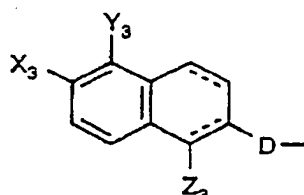
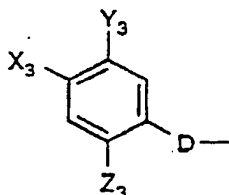
$4>m+n>0$ ,  $n<3$ , and  $m<3$ ;

25 or  $\text{R}_2$  is a group  $-\text{W}_5-(\text{C}_{1.5}$  alkyl or  $\text{C}_{2.5}$  alkenyl or  $\text{C}_{2.5}$  alkynyl) wherein  $\text{W}_5$  is a bond or is selected from  $-\text{O}-$ ,  $-\text{S}-$ ,  $-\text{SO}_2-$ ,

and  $\text{-NHC(O)-}$ , and the  $\text{C}_{1-5}$  alkyl,  $\text{C}_{2-5}$  alkenyl or  $\text{C}_{2-5}$  alkynyl moiety may be substituted with up to three groups selected independently from  $\text{OH}$ ,  $\text{-CONH}_2$ ,  $\text{-CSNH}_2$ ,  $\text{-CONHOH}$ ,  $\text{-CSNHOH}$ ,  $\text{-NHCHO}$ ,  $\text{-NHCONH}_2$ ,  $\text{-NHCSNH}_2$ ,  $\text{-NHSO}_2\text{NH}_2$  and  $\text{-SO}_2\text{NH}_2$ ;

- 5  $\text{-Z}_1\text{-R}_3$  is  $\text{-SO}_2(\text{OH})$ ,  $\text{-PO}(\text{OH})_2$ ,  $\text{-OSO}_2(\text{OH})$ ,  $\text{-NHSO}_2(\text{OH})$ ,  $\text{-NH-CO-COOH}$ ,  $\text{-SPO}(\text{OH})_2$ ,  $\text{-CH}_2\text{COOH}$ , tetrazol-5-yl or tetrazol-5-ylmethyl, or salts thereof;

or  $\text{Z}_1$  is  $\text{-O-}$ ,  $\text{-S-}$ ,  $\text{-NH-}$ , or  $\text{-CH}_2\text{-}$ , and  $\text{R}_3$  is a group of the formula:



- 10  $\text{D}$  is  $\text{-CH}_2\text{-}$ ,  $\text{-CO-}$ ,  $\text{-SO}_2\text{-}$ ,  $\text{-NH-SO}_2\text{-}$ ,  $\text{-NH-CO-}$ ,  $\text{-O-PO}(\text{OH})\text{-}$  or a salt thereof;

$\text{Z}_3$  is  $\text{H}$ ,  $\text{OH}$ ,  $\text{-CONH}_2$ ,  $\text{-CSNH}_2$ ,  $\text{-CONHOH}$ ,  $\text{-CSNHOH}$ ,  $\text{-NHCHO}$ ,  $\text{-NHCONH}_2$ ,  $\text{-NHCSNH}_2$ ,  $\text{-NHSO}_2\text{NH}_2$ ,  $\text{-SO}_2\text{NH}_2$ ,  $\text{-SO}_2(\text{OH})$ ,  $\text{-PO}(\text{OH})_2$ ,  $\text{-OSO}_2(\text{OH})$ ,  $\text{-NHSO}_2(\text{OH})$ ,  $\text{-COOH}$ , tetrazolyl-5-yl  
 15 or tetrazolyl-5-ylmethyl or a salt thereof,

with the proviso that when D is  $-\text{CH}_2-$ ,  $-\text{CO}-$ ,  $-\text{SO}_2-$ ,  $-\text{NHSO}_2-$  or  $-\text{NHCO}-$ , then  $\text{Z}_3$  is  $-\text{SO}_2(\text{OH})$ ,  $-\text{PO}(\text{OH})_2$ ,  $-\text{OSO}_2(\text{OH})$ ,  $-\text{NHSO}_2(\text{OH})$ ,  $-\text{COOH}$ , tetrazolyl-5-yl or tetrazolyl-5-ylmethyl or a salt thereof;

- 5  $\text{X}_3$  and  $\text{Y}_3$  independently are H,  $\text{NO}_2$ ,  $\text{SO}_2\text{NH}_2$ ,  $\text{CONH}_2$ ,  $\text{CF}_3$  or F; and

$\text{U}_4-\text{W}_4$  is  $-\text{CHCH}-$ ,  $-\text{CH}_2\text{CH}_2-$ ,  $-\text{C}(\text{OH})\text{CH}-$ ,  $-\text{CHC}(\text{OH})-$ ,  $-\text{CH}(\text{OH})\text{CH}_2-$ ,  $-\text{CH}_2\text{CH}(\text{OH})-$ ,  $-\text{CH}(\text{OH})\text{CH}(\text{OH})-$ ,  $-\text{C}(\text{O})\text{NH}-$ ,  $-\text{NHC}(\text{O})-$ ;

- 10  $\text{Y}_1$  is  $-\text{O}-$  or  $-\text{S}-$ ;

$\text{R}_4$  is H or, when  $\text{Y}_1$  is S,  $\text{S}(\text{CH}_2)_q\text{N}(\text{R}_9)_3^+$  and q is an integer 2-4, where  $\text{R}_9$  is H or  $\text{CH}_3$ ;

- $\text{R}_5$  is H;  $\text{C}_{1-6}$  alkyl,  $\text{C}_{2-6}$  alkenyl or  $\text{C}_{2-6}$  alkynyl, and the  $\text{C}_{1-6}$  alkyl,  $\text{C}_{2-6}$  alkenyl or  $\text{C}_{2-6}$  alkynyl moiety may be substituted  
 15 with OH,  $-\text{CONH}_2$ ,  $-\text{CSNH}_2$ ,  $-\text{CONHOH}$ ,  $-\text{CSNHOH}$ ,  $-\text{NHCHO}$ ,  $-\text{NHCONH}_2$ ,  $-\text{NHCSNH}_2$ ,  $-\text{NHSO}_2\text{NH}_2$  or  $-\text{SO}_2\text{NH}_2$ ; or aryl, aryl( $\text{C}_{1-2}$ )alkyl, heterocyclyl, or heterocyclyl( $\text{C}_{1-2}$ )alkyl which may optionally be substituted in the aryl or heterocyclyl moieties with one, two or three substituents selected independently from OH, F,  
 20 Cl,  $\text{NH}_2$ ,  $\text{CONH}_2$ ,  $\text{NHCOH}$ , and  $\text{SO}_2\text{NH}_2$ ;

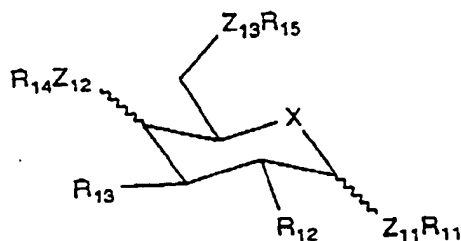
$\text{X}_1$  is  $-\text{O}-$ ,  $-\text{S}-$  or  $-\text{NH}-$ ;

$\text{R}_6$  is H or, when  $\text{X}_1$  is NH, acyl,  $\text{HOCNH-Val-Met-}$ ,  $\text{HOCNH-Ile-}$  (S,S)-dioxo-methionyl- or  $\text{HOCNH-Val-(pyran-4-on-2-yl)-alanyl-}$ ;

- 25 or a salt of such a compound.

Another family of substances called the hdo-family has also been synthesized. Hence, the invention also relates to novel compounds of the general formula

47



II

wherein

X is O, P, P(O), S, SO, SO<sub>2</sub>, CH<sub>2</sub>, C(OH)H, or a group NQ<sub>11</sub>, wherein Q<sub>11</sub> is H, OH, C<sub>1-24</sub> acyl or C<sub>1-24</sub> alkyl;

- 5 Z<sub>11</sub> is a bond, O, CH<sub>2</sub>, S, SO, SO<sub>2</sub>, or a group NQ<sub>12</sub>, wherein Q<sub>12</sub> is H, C<sub>1-24</sub> acyl or C<sub>1-24</sub> alkyl;

- R<sub>11</sub> is H; C<sub>1-24</sub> alkyl, C<sub>2-24</sub> alkenyl, or C<sub>2-24</sub> alkynyl, which may be substituted with one or more substituents independently selected from -OH, -COOH, -F, -Cl, -CONH<sub>2</sub>, -CSNH<sub>2</sub>,  
 10 -CONHOH, -CSNHOH, -NHCHO, -NHCONH<sub>2</sub>, -NHCSNH<sub>2</sub>, -NHSO<sub>3</sub>NH<sub>2</sub> and -SO<sub>2</sub>NH<sub>2</sub>; acyl; or -(CH<sub>2</sub>CH<sub>2</sub>O)<sub>s</sub>-H, wherein s = 1, 2, or 3;

- or R<sub>11</sub> is CH=CH-(CH<sub>2</sub>)<sub>n'</sub>-Q<sub>13</sub> or -(CH<sub>2</sub>)<sub>n'</sub>-Q<sub>13</sub>, wherein Q<sub>13</sub> is an aryl or a heteroaryl group substituted with -OH, -COOH, -F, -Cl, -CONH<sub>2</sub>, -CSNH<sub>2</sub>, -CONHOH, -CSNHOH, -NHCHO, -NHCONH<sub>2</sub>,  
 15 -NHCSNH<sub>2</sub>, -NHSO<sub>3</sub>NH<sub>2</sub> and -SO<sub>2</sub>NH<sub>2</sub>, and wherein n' ≥ 0;

- R<sub>12</sub> and R<sub>13</sub> are independently OH, H, F, Cl, OW<sub>11</sub>, or O(CO)W<sub>11</sub>, wherein W<sub>11</sub> is C<sub>1-24</sub> alkyl, C<sub>2-24</sub> alkenyl or C<sub>2-24</sub> alkynyl, or an aryl or a heteroaryl group substituted with -OH, -COOH, -F, -Cl, -CONH<sub>2</sub>, -CSNH<sub>2</sub>, -CONHOH, -CSNHOH, -NHCHO, -NHCONH<sub>2</sub>,  
 20 -NHCSNH<sub>2</sub>, -NHSO<sub>3</sub>NH<sub>2</sub> and -SO<sub>2</sub>NH<sub>2</sub>;

Z<sub>12</sub> is a bond, O, S, or CH<sub>2</sub>;

- R<sub>14</sub> is -(CH<sub>2</sub>)<sub>n''</sub>-Q<sub>14</sub>, wherein Q<sub>14</sub> is an aryl group or a heteroaryl group substituted with -OH, -COOH, -F, -Cl, -CONH<sub>2</sub>, -CSNH<sub>2</sub>, -CONHOH, -CSNHOH, -NHCHO, -NHCONH<sub>2</sub>, -NHCSNH<sub>2</sub>,  
 25 -NHSO<sub>3</sub>NH<sub>2</sub> and -SO<sub>2</sub>NH<sub>2</sub>, and wherein n'' = 0, 1, 2, or 3;

$Z_{13}$  is a bond, O,  $CH_2$ , S, SO,  $SO_2$ ,  $NQ_{14}Q_{15}$ , wherein  $Q_{14}$  is H,  $C_{1-24}$  acyl or  $C_{1-24}$  alkyl, and  $Q_{15}$  is CO or  $-C(O)W_{12}$ , wherein  $W_{12}$  is O or  $NW_{13}$ , wherein  $W_{13}$  is H, OH,  $C_{1-24}$  acyl or  $C_{1-24}$  alkyl;

- 5  $R_{15}$  is H;  $C_{1-24}$  alkyl,  $C_{2-24}$  alkenyl or  $C_{2-24}$  alkynyl, which alkyl, alkenyl and alkynyl may be substituted with one or more substituents independently selected from, -OH, -COOH, -F, -Cl, -CONH<sub>2</sub>, -CSNH<sub>2</sub>, -CONHOH, -CSNHOH, -NHCHO, -NHCONH<sub>2</sub>, -NHCSNH<sub>2</sub>, -NHSO<sub>3</sub>NH<sub>2</sub> and -SO<sub>2</sub>NH<sub>2</sub>; acyl or  $-(CH_2CH_2O)_s-H$ , where-  
10 in  $s = 1, 2, \text{ or } 3$ ;

or  $R_{15}$  is  $CH=CH-(CH_2)_{n'}-Q_{13}$ , or  $-(CH_2)_{n'}-Q_{13}$ , wherein  $Q_3$  is as defined above and wherein  $n' \geq 0$ ;

or a salt of such a compound.

- 15 In the present context, the terms " $C_{1-5}$ ,  $C_{1-6}$  and  $C_{1-24}$  alkyl" is intended to mean alkyl groups with 1-5, 1-6 and 1-24 carbon atoms, respectively, which may be straight or branched or cyclic such as methyl, ethyl, propyl, isopropyl, butyl, isobutyl, tert.butyl, hexyl, octyl, dodecyl, cyclopentyl, cyclohexyl, etc.

- 20 Further, as used herein, the terms " $C_{2-5}$ ,  $C_{2-6}$  and  $C_{2-24}$  alkenyl" is intended to mean mono- or polyunsaturated alkyl groups with 2-5 and 2-24 carbon atoms, respectively, which may be straight or branched or cyclic in which the double  
25 bond(s) may be present anywhere in the chain or the ring, for example vinyl, 1-propenyl, 2-propenyl, hexenyl, decenyl, 1,3-heptadienyl, cyclohexenyl etc. Some of the substituents exist both in a cis and a trans configuration. The scope of this inventions comprises both the cis and trans forms.

- 30 In the present context, the terms " $C_{2-5}$ ,  $C_{2-6}$  and  $C_{2-24}$  alkynyl" is intended to mean a straight or branched alkyl group with 2-5 and 2-24 carbon atoms, respectively, and incorporat-

ing one or more triple bond(s), e.g. ethynyl, 1-propynyl, 2-propynyl, 2-butyryl, 1,6-heptadiynyl, etc.

The terms "C<sub>1-6</sub> and C<sub>1-24</sub> alkoxy" designate alkyl groups as defined above comprising an oxy function.

- 5 In the present context, the term "aryl" is intended to mean phenyl and naphthyl. The term "heteroaryl" is intended to mean a cyclic aromatic system, wherein at least one non-carbon atom contributes to the  $\pi$  bonding system.

Examples of substituted aryl groups are: 3-nitrophenyl, 3-  
10 hydroxyphenyl, 4-hydroxyphenyl, 3,4-dihydroxyphenyl, 3-carboxamidophenyl, 3-formamidylphenyl, 3-acetamidylphenyl, 3-fluoronaphth-2-yl, 7-fluoronaphthyl, 3,7-difluoronaphthyl, 3-hydroxynaphthyl, 7-hydroxynaphthyl, 3,7-dihydroxynaphthyl, 3-fluoro-7-hydroxynaphthyl, 7-fluoro-3-hydroxynaphthyl, 4-  
15 fluoronaphth-2-yl, 6-fluoronaphth-yl, 8-fluoronaphth-2-yl, 4,6-difluoronaphth-2-yl etc.

Examples of heterocyclic and heteroaryl groups are pyrrolyl, furanyl, 2,3-dihydrofuranyl, tetrahydrofuranyl, thienyl, 2,3-dihydrothienyl, tetrahydrothienyl, imidazolyl, oxazolyl,  
20 thiazolyl, pyrazolyl, indolyl, pyrazinyl, dioxolanyl, dioxan-yl, 1,3,5-trioxanyl, tetrahydrothiapyranyl, dithiolanyl, pyrazolidinyl, iminazolidinyl, sym-triazinyl, sym-tetrazinyl, quinazolinyl, pteridinyl, isoindolyl, 1,2,4-triazolyl, 1,2,3-triazolyl, benzimidazolyl, indazolyl, benzofuranyl, isoben-  
25 zofuranyl, benzothiophenyl, thienothiophenyl, isoxazolyl, 1,2,5-oxadiazolyl, isothiazolyl, 1,3,4-thiadiazolyl, benzoxazolyl, benzothiazolyl, azaindolyl, oxoindolyl, hydroxyindolyl, N-oxyisoquinolyl etc.

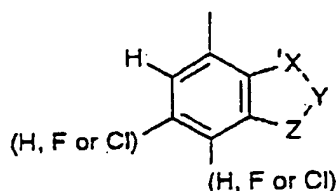
In the present context, the term "acyl" (e.g. C<sub>1-24</sub> acyl) is  
30 intended to designate the acyl residue of a carboxylic acid or a sulphonic acid moiety comprising a carbonyl or sulphonyl group and an organic moiety. Examples of acyl groups include C<sub>1-24</sub> alkanoyl (e.g. formyl, acetyl, propionyl, butyryl,

isobutyryl, valeryl, isovaleryl, pivaloyl and hexanoyl),  
C<sub>1-24</sub> alkenoyl (e.g. acryloyl, metacryloyl, crotonoyl, iso-  
crotonoyl, 2-pentenoyl, 3-pentenoyl, 2-methylpentenoyl, 3-  
pentenoyl, 3-phenylpropenoyl, 2-phenyl-trans-propenoyl, 2,4-  
5 hexadienoyl), C<sub>1-24</sub> alkynoyl (e.g. propyonyl, 2-butyneoyl, 3-  
butyneoyl, 2-methyl-3-butyneoyl, 2,2-dimethylbutyneoyl, 2-penty-  
noyl, 3-pentyneoyl, 2-pentyne-4-trans-enoyl), C<sub>1-24</sub> alkoxycar-  
bonyl (e.g. methoxycarbonyl, ethoxycarbonyl, propoxycarbonyl,  
butoxycarbonyl and t-butoxycarbonyl), C<sub>1-24</sub> alkenyloxycarbon-  
10 yl (e.g. cis-2-butenyloxycarbonyl, 1-methyl-2-propenyloxycar-  
bonyl, 1,1-dimethyl-2-propenyloxycarbonyl, trans-2-butenyl-  
oxycarbonyl), aroyl (e.g. benzoyl), heterocyclylcarbonyl  
(e.g. 2-furoyl, 3-furoyl, 2-furanoyl, 3-furanoyl, 2-pyrrol-  
carboxyl, 3-pyrrolcarboxyl, 2-thenoyl, 3-thenoyl, 2-indolcar-  
15 boxyl, 3-indolcarboxyl, 1-naphthanoyl and 2-naphthanoyl),  
etc.

The term "salt" is intended to comprise a salt such as an  
organic acid addition salt (e.g. acetate, valerate,  
salicylate, galacturonate, gluconate, tannate, trifluoroace-  
20 tate, maleate, tartrate, methanesulfonate, benzenesulfonate,  
formate, thiocyanate and toluenesulfonate), an inorganic  
acid addition salt (e.g. hydrochloride, hydrobromide, hydro-  
iodide, dihydrochloride, dihydrobromide, dihydroiodide,  
sulphate, hydrogensulphate, halosulphate such as iodosul-  
25 phate, nitrate, phosphate, and carbonate) or a salt with an  
amino acid (e.g. arginine, aspartic acid and glutamic acid)  
or a metal salt such as an alkali metal salt (e.g. sodium  
salt and potassium salt) and an earth alkali metal salt (e.g.  
calcium salt and magnesium salt), an ammonium salt, an  
30 organic alkali salt (e.g. trimethylamine salt, triethylamine  
salt, pyridine salt, picoline salt, dicyclohexylamine salt  
and N,N'-dibenzylethylenediamine salt), and hydrates thereof.

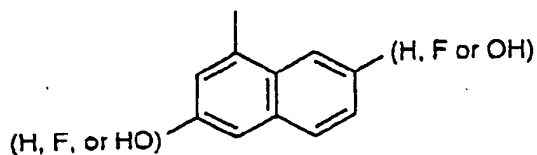
When the substituent R<sub>5</sub> designates heterocyclyl, it is pre-  
ferred that the substituent designates a heterocyclyl group  
35 of the formula



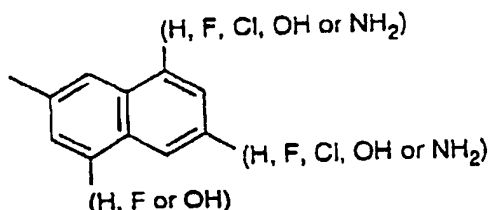


wherein 'X' is -CH-, -CH<sub>2</sub>-, -O-, -N-, -S-, -S=O, -N=O or -CO-,  
 Y is -CH- or -NH-, and Z is -CH-, -CH<sub>2</sub>-, -O-, -S-, -N-, -CO-,  
 -S=O or -N=O, especially a group selected from the group  
 consisting of inden-7-yl, benzofuran-4-yl, isobenzofuran-4-  
 5 yl, thionaphthen-4-yl, isothionaphthen-4-yl, 2-oxo-inden-7-  
 yl, 2-oxo-inden-4-yl, inden-4-yl, benzofuran-7-yl, isoben-  
 zofuran-7-yl, thionaphthen-7-yl, isothionaphthen-7-yl, 1-oxo-  
 thionaphthen-4-yl, 1-oxo-thionaphthen-7-yl, anthran-4-yl,  
 anthran-7-yl, thioanthran-4-yl, thioanthran-7-yl, benzthio-  
 10 zol-4-yl, benzthiozol-7-yl, 2H-2-isobenzo-1,3-dion-7-yl,  
 isobenzofuran-5-yl, isobenzofuran-6-yl, 3H-2-oxo-benzofuran-  
 5-yl, 3H-2-oxo-benzofuran-6-yl, 3H-2-oxothionaphthen-5-yl,  
 3H-2-oxothionaphthen-6-yl, indol-5-yl, indol-6-yl, 3H-2-  
 oxoindol-5-yl, 3H-2-oxoindol-6-yl, 3H-2-oxobenzoxazol-5-yl,  
 15 3H-2-oxobenzoxazol-6-yl, benzothiazol-5-yl, benzothiazol-6-  
 yl, 2-oxobenzo-1,3-dithiol-5-yl, 2-oxobenzo-1,3-dithiol-6-yl,  
 3H-2-oxobenzimidazol-5-yl, 3H-2-oxobenzimidazol-6-yl, ben-  
 zoxathiol-5-yl, benzoxathiol-6-yl, 3H-2-oxobenzthiazol-5-yl  
 and 3H-2-oxobenzthiozol-6-yl.

20 It is also preferred that the substituent R<sub>5</sub> is a group of  
 the formula



or that  $R_5$  is a group of the formula

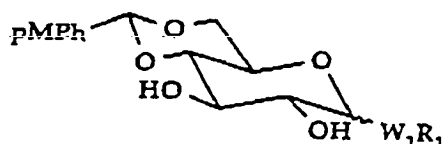


With respect to the substituent  $R_2$  being heterocyclcyl, it is especially preferred that it is selected from the group consisting of isobenzofuran-5-yl, isobenzofuran-6-yl, 3H-2-oxo-benzofuran-5-yl, 3H-2-oxo-benzofuran-6-yl, 3H-2-oxothio-  
 5 naphthen-5-yl, 3H-2-oxothionaphthen-6-yl, indol-5-yl, indol-6-yl, 3H-2-oxoindol-5-yl, 3H-2-oxoindol-6-yl, 3H-2-oxoben-  
 zoxazol-5-yl, 3H-2-oxobenzoxazol-6-yl, benzothiazol-5-yl, benzothiazol-6-yl, 2-oxobenzo-1,3-dithiol-5-yl, 2-oxobenzo-  
 10 1,3-dithiol-6-yl, 3H-2-oxobenzimidazol-5-yl, 3H-2-oxobenzimidazol-6-yl, benzoxathiol-5-yl, benzoxathiol-6-yl, 3H-2-oxobenzthiazol-5-yl and 3H-2-oxobenzthiozol-6-yl.

The exact number of substituents present on an alkyl, alkenyl or alkynyl moiety  $R_1$  will be dependent on the length of the  
 15 carbon chain, the purpose of the substituents being to cause the entire group  $R_1$  to be compatible with water since, in the chaperone-ligand-complex such as the PapD-ligand complex, the group  $R_1$  will extend into the surrounding aqueous medium.  
 Thus, for a fairly short carbon chain such as up to four  
 20 carbon atoms, it is contemplated that one of the above polar

substituents will be sufficient, in particular when the substituent is located terminally whereas, for longer chains, a larger number of substituents, such as a substituent for every other carbon atom, may be required.

- 5 The 4,6-O-(4'-Methoxy)phenylmethyldiene- $\alpha$ -D-glucohexopyranoside or 4,6-O-(4'-Methoxy)phenylmethyldiene- $\beta$ -D-glucohexopyranoside glycosides:



A

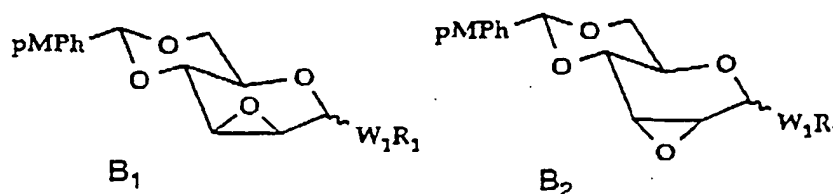
- (used here as preferred examples, but other arylmethyldiene or vinylidene acetals may be used) can be prepared as follows: Peracylated glucose is reacted with, e.g. hydrogen bromide or hydrogen chloride in a suitable solvent such as, e.g. acetic acid or dichloromethane, to form per-O-acylated glycosyl bromide or chloride (O-acylation and glycosyl halide synthesis: M. L. Wolfrom and A. Thompson, *Methods in Carbohydrate Chemistry*, Vol. 2, 211-215, ed. by R.L. Whistler and L. Wolfrom, Academic Press, New York, 1963; G.Hewit and G. Fletcher Jr., *ibid*, 226-228; and R.U. Lemieux, *ibid*, 223-224)

- The suitably protected, when necessary, aglycone alcohol or thiol ( $H-W_1R_1-PG_1$  or  $H-W_1R_1$ ) (protecting groups: *Protective Groups in Organic Synthesis*, Editors T. W. Greene and P. G. M. Wuts, John Wiley & Sons, Inc., New York, 1991) is reacted with the per-O-acylated glucose using a Lewis acid such as borontrifluoride etherate (R. J. Ferrier and R. H. Furneaux, *Carbohydr. Res.*, 52 (1976) 63-68; J. Dahmén, T. Frejd, G. Grönberg, T. Lave, G. Magnusson, and G. Noori, *Carbohydr. Res.* 116 (1983) 303-307) or trimethylsilyltrifluoromethanesulphonate (T. Ogawa, K. Beppu, S. Nakabayashi, *Carbohydr. Res.*, 93 (1981) C6-C9) as promoters. The reaction is carried out in a

suitable solvent such as chloroform, dichlorometane, or toluene. When the monosaccharide derivative in question is a per-O-acylated glycosyl bromide or chloride, promoters such as silver trifluoromethane sulphonate or mercury (II)salts  
5 (H. Paulsen, *Angew. Chem. Int. Ed. Engl.*, 21 (1982) 155-173) can be used and the reactions are carried out in a suitable solvent such as dichlorometane or toluene. The glucose  $W_1R_1$  or  $W_1R_1PG_1$ -glycosides are obtained after de-O-acylation using sodium methoxide (A. Thompson, M. L. Wolfrom, and E. Pascu,  
10 page 215-220, *Methods in Carbohydrate Chemistry*, Vol II, Editors: R. L. Whistler and M. L., E. Wolfrom, Academic Press, New York, 1963) in methanol or in methanol containing a co-solvent such as dichloromethane or tetrahydrofuran.

The 4,6-(4'-methoxy)benzylidene acetals are then obtained by  
15 reaction with 4-methoxybenzaldehyde dimethyl acetal and acid in a polar non-protic solvent such as e.g. dimethyl formamide, acetonitrile or tetrahydrofuran (J.J. Patroni et al., *Aust. J. Chem.* 1988, (41), 91-102; for other methods of acetal formation, see for example A. N. de Belder, 1979, *adv. Carbo-*  
20 *hydr. Chem. Biochem.*, 34, 179 and references cited therein).

The epoxides B1 or B2

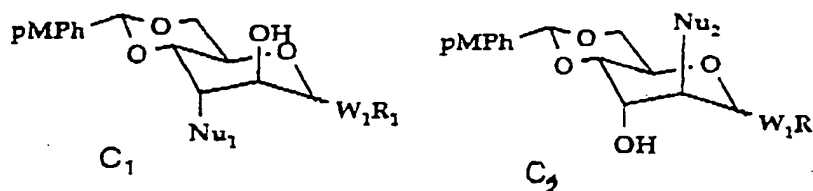


are then obtained through sulphonate esters: The *manno* epoxides B1 can be prepared by reacting the glucoside derivative A with sodium hydride and p-toluenesulphonylimidazol in  
25 dimethylformamide (D. Hicks and B. Fraser-Reid; *Synthesis* 1974, 203) or with sodium hydride and p-toluenesulphonyl chloride in tetrahydrofuran (V.S. Murthy et al, *Synthetic Commun.* 1993, 23(3), 285-289).

The allo epoxides B2 can be prepared by reacting the glucoside derivative A with methylsulphonyl or p-toluenesulphonyl chloride in pyridine and treating the resulting methylsulphonate diester with sodium ethoxide in ethanol (Y.

- 5 Ali, A. C. Richardson, *Carbohydrate Res.* 1967, 5, 441-448; N. Richtmeyer, *Methods in Carbohydrate Chemistry*, Vol 1,107).

The epoxides B1 or B2 can be reacted with suitable nucleophilic reagents to yield the diaxially substituted allo hexopyranosides C1 and C2



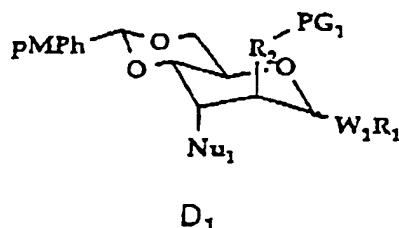
- 10 (for general references on the use of epoxides, see e.g. J. Gorzynski Smith, *Synthesis*, 1984, 8, 629-656 Masamune S., Choy W., Petersen J S, and Sita L R, *Angew. Chem. Int. Ed. Engl.*, 1985, 24, 1-76; A. S. Rao et al., *Tetrahedron Lett.*, 1983, 39, 2323).
- 15 When the first atom of  $R_2$  and  $Z_1$  (as defined above) connected to the carbohydrate moiety in the desired final product is a nitrogen (=the nucleophilic atom), then the preferred nucleophile  $Nu_1$  or  $Nu_2$  is azide ( $N_3^-$ ). The epoxide is treated with sodium azide and ammonium chloride in boiling 2-methoxy-
- 20 ethanol (R. D. Guthrie and D. Murphy; *J. Chem. Soc.* 1963, 5288-5294).

- When the nucleophilic atom is oxygen or sulphur, the preferred general method of epoxide opening involves treatment with suitably protected alcohol or thiol in the presence of
- 25 neutral alumina in ether (G.H. Posner and D.Z. Rogers, *J. Am. Chem. Soc.* 1977, 99, 8208; G.H. Posner, D.Z. Rogers and A.

Romero, *Isr. J. Chem.* 1979, 18, 259; and G.H. Posner, M. Hulce and R.K. Rose, *Synth. Commun.* 1981, 11, 737).

When the nucleophilic atom is carbon, the most commonly used reagents are organomagnesium, organolithium, organocopper, organoaluminium and organoboron compounds (J. Gorzynski Smith, *Synthesis*, 1984, 8, 629-656 and references cited therein).

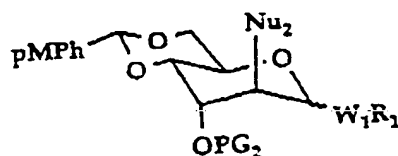
When the product is an allo hexopyranoside C1, the 2-hydroxy function can either be blocked with a protective group that allows for the introduction of the R<sub>2</sub> functionality at a later stage (preferred if R<sub>2</sub> is an ester, not shown in figure), or the suitably protected, when necessary, functionality R<sub>2</sub> is introduced to produce D1



For example, OH-groups to ethers or esters (*Protective Groups in Organic Synthesis*, Editors T. W. Greene and P. G. M. Wuts, John Wiley & Sons, Inc., New York, 1991); OH-groups to carbonates (J. March, *Advanced Organic Chemistry-Reaction Mechanisms, and Structure*, 3rd Edn. John Wiley & Sons, New York, 347 (1985), and references cited therein); OH-groups to carbamates (J. March, *Advanced Organic Chemistry-Reaction Mechanisms, and Structure*, 3rd Edn. John Wiley & Sons, New York, 791-792 (1985), and references cited therein); OH-groups to alkylgroups via exomethylene derivatives and subsequent hydrogenation or via other routes (H. O. H. House, *Modern Synthetic Reactions*, 2nd Edn. W. A Benjamin, Inc., Menlo Park, C.A., 1-130 (1972), and references cited therein; J. Yoshimura, *Adv. Carbohydr. Chem. Biochem.*, 42 (1984) 69-134); and exchange of OH-groups for heterocyclic groups, via

different routes (A. R. Katritzky, *Handbook of Heterocyclic Chemistry*, Pergamon Press, Oxford, 1985).

When the product is an *allo* hexopyranoside C2, the 3-hydroxy function is blocked with a protective group that allows for the introduction of the Z<sub>1</sub>-R<sub>3</sub> functionality at a later stage resulting in intermediates of the type D2

D<sub>2</sub>

(*Protective Groups in Organic Synthesis*, Editors T. W. Greene and P. G. M. Wuts, John Wiley & Sons, Inc., New York, 1991) or transformed to a *manno* hexopyranoside intermediate D3

D<sub>3</sub>

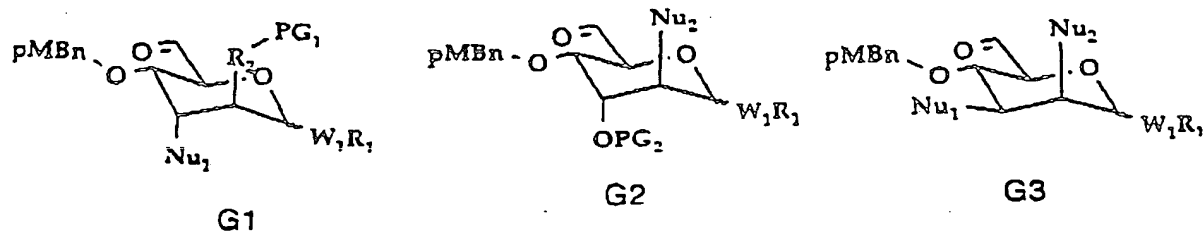
- 10 where Nu<sub>1</sub> is a protected or masked form of the functionality Z<sub>1</sub>. OH-groups to ethers or esters (*Protective Groups in Organic Synthesis*, Editors T. W. Greene and P. G. M. Wuts, John Wiley & Sons, Inc., New York, 1991); OH-groups to azido-
- 15 groups: J. March, *Advanced Organic Chemistry-Reaction Mechanisms, and Structure*, 3rd Edn. John Wiley & Sons, New York, 380, (1985), and references cited therein; H. H. Baer, *Pure Appl. Chem.*, 61(7) (1989) 1217-1234, and references cited therein; OH-groups to aminogroups via azides or other routes
- 20 (*March, Advanced Organic Chemistry-Reaction Mechanisms, and Structure*, 3rd Edn. John Wiley & Sons, New York, 1106, 798-800 (1985), and references cited therein;

H. H. Baer, *Pure Appl. Chem.*, 61(7) (1989) 1217-1234, and references cited therein).

The 4,6-O-acetal function is then reductively opened to yield either the function  $R_6$  in the case when  $R_6$  is an ether, or the intermediates F1, F2 or F3 with a hydroxy function on position 6 (reductive opening of acetals, see Garegg P J and Hultberg H, *Carbohydr. Res.* 1981, 93, c10-11; Garegg P J, Hultberg H, and Wallin S, *Carbohydr. Res.* 1982, 108, 97-101; Liptak A, Jodal I, Nanasi P, *Carbohydr. Res.* 1975, 44, 1-11; Baker D C, Horton D, Tindall C G, *Methods in Carbohydr. Chem.*, 1976 Vol 6, 3-6; Mikami T, Asano H, Mitsunobu O, *Chem. Lett.* 1987, 10, 2033-2036; Ek M, Garegg P J, Hultberg H, Oscarsson S, *J. Carbohydr. Chem.* 1983, 2, 305-311; Hunter R, Bartels B, Michael J P, *Tetrahedron Lett.* 1991, 32, 1095-1098; Rao S P, Grindley T B, *Carbohydr. Res.* 1991, 218, 83-93; Hunter, R. Bartels, B. *J. Chem. Soc. Chem. Commun.* 1991, 2887-2888).

For example, the intermediates of type D1, D2 or D3 are treated with sodium cyanoborohydride and chlorotrimethylsilane in acetonitrile, (R. Johansson and B. Samuelsson, *J. Chem. Soc. Perkin Trans. 1*, 1984, 2371-2374) or borane-trimethylamine complex and aluminium trichloride. The regiochemical outcome of the reaction is often solvent-dependent (Ek M, Garegg P J, Hultberg H, Oscarsson S, *J. Carbohydr. Chem.* 1983, 2, 305-311).

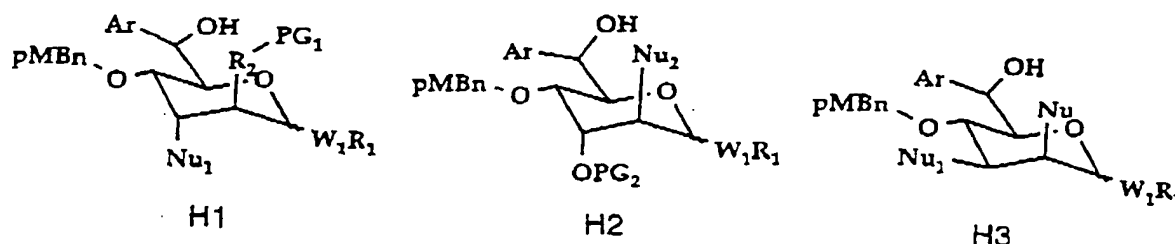
The aldehyde intermediates of type G1, G2 or G3





are obtained by oxidation of the corresponding 6-alcohol intermediates of type D1, D2 or D3, preferably by the Swern procedure. (Mancuso A J, Swern D, *Synthesis*, 1981, 165-185; Tidwell T, *Synthesis*, 1990, 857-870; for other oxidation methods, see A. H. Haines, 1988, *Methods for the Oxidation of Organic Compounds*, Chapter 2, Academic Press, San Diego, and references cited therein).

In the next step a carbon nucleophile is added to the aldehyde function of the intermediates of type G1, G2 or G3. Preferably, a suitably protected alkyllithium or aryllithium reagent or a grignard reagent is added to the aldehyde in an ether or hydrocarbon solvent to produce the secondary alcohols H1, H2 and H3:



For reactions of aldehydes with organolithium and organomagnesium compounds, see J. March, *Advanced Organic Chemistry-Reaction Mechanisms, and Structure*, 3rd Edn. John Wiley & Sons, New York, 347 (1985), and references cited therein.

For reactions of aldehydes with organolithium and organomagnesium and other carbon nucleophils, see Evans D A, *Aldrichim. Acta*, 1982, 15, 23, and references cited therein.

For specific examples on the use and preparation of arylsubstituted phenyllithium and grignard reagents, see Ames M M, Castagnoli Jr. N, *J. Labelled Compd.*, 1974, 10(2), 195-205; DE 3807910 A1; Mills R J, Snieckus V, *Polynucl. Aromat. Hydrocarbons*, [Pap. Int. Symp.], 8th, Meeting 1983, 913-24. Edited by: Cooke M and Dennis A J, Battelle Press 1985: Columbus,

- Ohio; Iriye R, Furukawa K, Nishida R, Kim C, Fukami H, *Bio-sci. Biotechnol. Biochem.* 1992, 56(11), 1773-5; Comber M F, Sargent M V U, *J. Chem. Soc., Chem. Commun.*, 1991, (3), 190-2; Hirai T, Yoshizawa A, Nishiyama I, Fukumasa M, Shiratori N, Yokoyama A, EP 341686 A2; Leeson P D, Emmett J C, Shah V P, Showell G A, Novelli R, Prain H D, Benson M G, Ellis D, Pearce N J, Underwood A H, *J. Med. Chem.* 1989, 32(2), 320-36; Meltzer P C, Liang A Y, Brownell A L, Elmaleh D R, Madras B K, *J. Med. Chem.*, 36(7), 855-62; Willard A K, Novello F C, Hoffman W F, Cragoe Jr E J. US 4,459,423.

For substituted phenyllithium reagents that can be further elaborated into heterocyclic compounds, see: Lang H J, Muschaweck R, AU 514406 B2; Lang H J, Muschaweck R, Hropot M, HU 19761; Lang H J, Muschaweck R, Hropot M, DE 2737195.

- 15 For heteroaromatic aryllithiums and grignard reagents, see: Yang Y, Martin A R, Nelson D L, Regan J, *Heterocycles*, 1992, 34(6), 1169-75.

For examples of other organometallic reagents for the stereoselective synthesis of secondary alcohols from aldehydes:

- 20 Homoallyl alcohols with crotylmolybdenum complexes:  
Faller J W, John J A, Mazziere M R, *Tetrahedron Lett.* 1989, 30, 1769-1772.

- Homoallyl alcohols with titanium complexes: Riediker M, Duthaler R O, *Angew. Chem. Int. Ed. Engl.*, 1989, 28, 494-495.

25 3-Pyrollyl alcohols with silyl protected 3-lithiopyrrole:  
Bray B L, Mathies P H, Naef R, Solas D R, Tidwell T T, Artis D R, Muchowski J M, *J. Org. Chem.*, 1990, 55, 6317-6328.

Allylic alcohols with E-vinylalane: A P Kozikowski and Jiang-Ping Wu, *Tetrahedron Lett.* 1990, 30, 4309-4312 and references cited therein.

5 Trans-allylic diols with vinylstannanes: Corey E J, Wollenberg R H, *J. Org. Chem.*, 1975, 40, 2265-2266.

Pyrrolidine carbinols with  $\alpha$ -lithio pyrrolidine amidines: Sanner M A, *Tetrahedron Lett.* 1989, 30, 1909-1912.

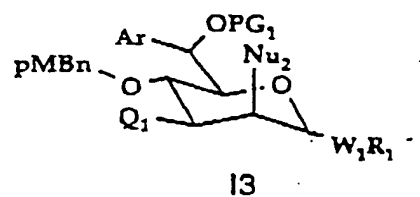
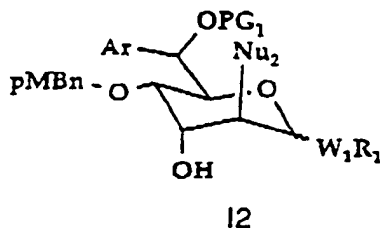
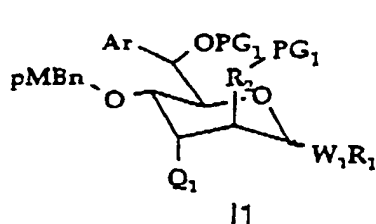
Organozinc reagents: Fürstner A, *Synthesis*, 1989, 571-590, and references cited therein.

10 Thereafter, the substituent on the 3-position ( $\text{Nu}_1$  or  $\text{OPG}_1$ ) is transformed into a nucleophile in order to install the negatively charged functionality  $\text{Z}_1\text{-D-R}_3$ .

The secondary 6-alcohols H1, H2 and H3 are first protected, (Protective Groups in Organic Synthesis, Editors T. W. Greene  
15 and P. G. M. Wuts, John Wiley & Sons, Inc., New York, 1991) or transformed to amino functions in the cases where  $\text{X}_1\text{-R}_6$  forms a peptide functionality (OH-groups to aminogroups via azides or other routes: See for examples March, *Advanced Organic Chemistry-Reaction Mechanisms, and Structure*, 3rd  
20 Edn. John Wiley & Sons, New York, 1106, 798-800 (1985), and references cited therein; H. H. Baer, *Pure Appl. Chem.*, 61(7) (1989) 1217-1234, and references cited therein).

For peptide synthesis, see Gross and Meienhofer, *The Peptides*, 3 vol., Academic Press, New York, 1979-1981; Grant G A  
25 et al., *Synthetic Peptides, A Users Guide*, 1992, W. A. Freeman and Company, New York, and references cited therein).

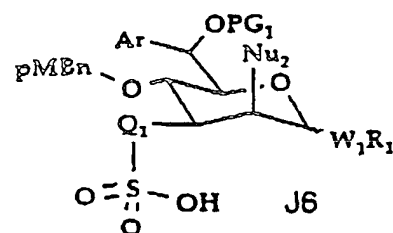
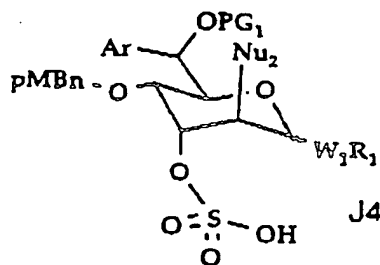
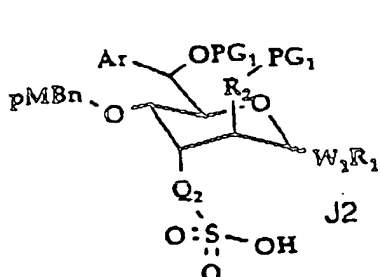
The 3-position is deprotected to an alcohol, thiol or amine intermediate I1, I2 or I3



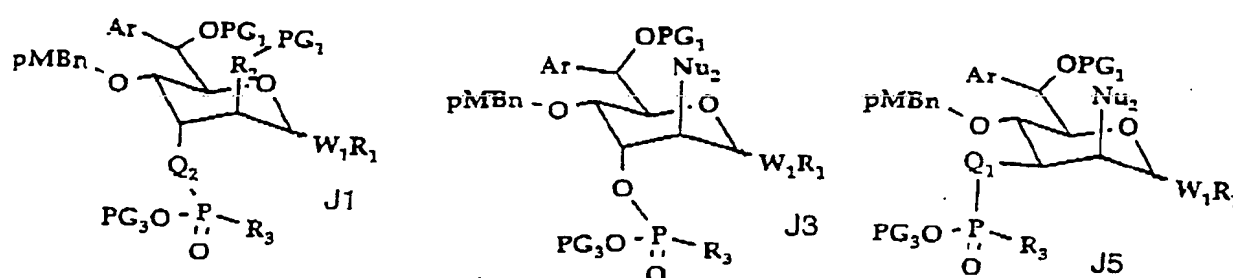
(*Protective Groups in Organic Synthesis*, Editors T. W. Greene and P. G. M. Wuts, John Wiley & Sons, Inc., New York, 1991, and references cited therein).

For example, treatment of I1, where R<sub>2</sub>-PG<sub>1</sub> is a combination  
 5 of ether functions and Q<sub>1</sub> is azide, with gaseous hydrogen sulphide in pyridine and water (T. Adashi, Y. Yamada, I. Inoue and M. Saneyoshi, *Synthesis*, 1977, 45). For other methods of azide reductions, see Poopeiko N E, Pricota T I, Mikhailopulo I A, *Synlett*, 1991, 5, 342; Samano M C, Robins M  
 10 J, *Tetrahedron Lett.* 1991, 32, 6293-6296; Rakotomanomana N, Lacombe J-M, Pavia A, *Carbohydr. Res.*, 1990, 197, 318-323; Malik A A, Preston S B, Archibald T G, Cohen M P, Baum K, *Synthesis*, 1989, 450-451; Maiti S, Spevak P, Reddy N, *Synt. Commun.* 1988, 18, 1201-1206; Bayley H, Standring D N, Knowles  
 15 J R, *Tetrahedron Lett.*, 1978, 39, 3633-3634).

The intermediates I1, I2 or I3 are sulphated with sulphur trioxide-pyridine or -triethylamine complex to obtain the intermediates J2, J4 and J6



(see, for example J. Basten, G. Jaurand, B. Olde-Hanter, M. Petitou and C.A.A. van Boeckel, *Bioorg. Med. Chem. Lett.*, 1992, 2(9), 901-904 and references cited therein; J. Basten, G. Jaurand, B. Olde-Hanter, P. Duchaussoy, M. Petitou and C.A.A. van Boeckel, *Bioorg. Med. Chem. Lett.*, 1992, 2(9), 905-910, and references cited therein; Böcker, T. Lindhorst, TK. Thiem, J. Vill, V. *Carbohydr. Res.* 1992, 230, 245-256) or coupled to form the phosphoester intermediates J1, J3 and J5



Since the phosphorus-nitrogen bond is known to be acid labile, the intermediates leading to phosphodiester end-products are preferred (M. Selim and T. N. Thanh, *C.R. Seances Acad. Sci.*, 1960, 250, 2377).

For example, the alcohol intermediates I1, I2 or I3 are treated with 2,2,2-trichloroethyl 2-chlorophenyl phosphochloridate in chloroform and pyridine to form the phosphate triesters. The 2,2,2-trichloroethyl group is removed by treatment with zinc powder and the resulting phosphate diester is activated with 3-nitro-1-(2,4,6-triisopropylbenzenesulphonyl)-1,2,4-triazole and coupled with the alcohol R<sub>3</sub>-OH to form the intermediates J1, J3 and J5. The 2-chlorophenyl group is removed by treatment with pyridine-2-aldoxime and N,N,N,N-tetramethylguanidine in moist pyridine (see for example P. J. Garegg, R. Johansson, I. Lindh and B. Samuelsson, *Carbohydr. Res.* 1986, 150, 285-289).

Alternatively, by the phosphite triester approach, alcohol intermediates I1, I2 or I3 are treated with phenyl chloro-N,N-diisopropylphosphoramidite in acetonitrile to form carbo-

hydrate phosphoamidites. After purification by chromatography, these are exposed to an alcohol  $R_3$ -OH in the presence of a mild acid, such as pyridinium p-toluenesulphonate and treated with tert-butyl hydroperoxide to form phosphate diesters J1, J3 and J5 ( $PG_3 = H$ ) (H-N. Caro, M. Martín-Lomas and M. Bernabé, *Carbohydr. Res.* 1993, 240, 119-131, and references cited therein).

For a review on phosphodiester in DNA synthesis, see Narang S, *Tetrahedron*, 1983, 39, 1-22 and D. W. Hutchinson, 1991, Chapter 3 in *Chemistry of Nucleosides and Nucleotides*, vol 2, ed. B. Townsend, Plenum Press, New York, and references cited therein. For other examples on carbohydrate phosphodiester synthesis, see for example Ichikawa Y, Sim M M, Wong C H, *J. Org. Chem.* 1992, 57, 2943-2946 and Schmidt R R, Braun H, Jung K-H, *Tetrahedron Lett.* 1992, 33, 1585-1588. For synthesis of modified phosphodiester linkages, See R. S. Varma, 1993, *Synlett*, 621-636, and references cited therein.

To obtain arylphosphonic acid esters and amides J1, J3 and J5, where  $R_3$  is an alkyl or aromatic group the arylphosphonic acid  $R_3$ - $PO_3H_2$  is coupled to the alcohol intermediates I1, I2 or I3, for example, with a carbodiimide reagent, or treatment of the phosphonic dichlorides with the alcohol intermediates I1, I2 or I3 in pyridine (T. H. Siddal III, and C. A. Prohas-ka, *J. Am. Chem.* 1962, 84, 3467).

The alkylphosphonic triesters are formed from trialkylphosphites and alkyl halides by the Arbuzov reaction (Arbuzov, *Pure Appl. Chem.* 1964, 9, 307-335, J. March, *Advanced Organic Chemistry-Reaction Mechanisms, and Structure*, 3rd Edn. John Wiley & Sons, New York, 347 (1985), and references cited therein).

For the preparation of arylphosphonic triesters via phosphorus trichloride, see for example: K. Sasse, *Methoden der Organischen Chemie* (Houben-Weyl), 4th ed., Vol. 12/1, Georg Thieme, Stuttgart, 1963, p. 314 and p. 392 and references

cited therein; G. M. Kosolapoff, *Org. React.* 6, 273 (1951), and references cited therein; L.D. Freedman and G. O. Doak, *Chem. Rev.* 1957, 57, 479, and references cited therein.

For the preparation of arylphosphonic triesters via organo-  
5 magnesium or organolithium reagents, see for example: K. Sasse, *Methoden der Organischen Chemie* (Houben-Weyl), 4th ed., Vol. 12/1, Georg Thieme, Stuttgart, 1963, p. 372, and references cited therein; G. M. Kosolapoff, *Org. React.* 6, 273 (1951), and references cited therein.

- 10 Intermediates J1, J2, J3, J4, J5 and J6 are deprotected to form the final products (*Protective Groups in Organic Synthesis*, Editors T. W. Greene and P. G. M. Wuts, John Wiley & Sons, Inc., New York, 1991, and references cited therein) and transformed to their sodium or potassium salts.
- 15 It is justified to assume that the compounds described above are capable of interacting with sites in PapD and other periplasmic chaperones. In order to establish that this is really the case, assays like those described in the examples should be performed.
- 20 Thus, a preferred compound of the invention is a compound as described above, which causes a prevention, inhibition or enhancement of the binding of G1'-19'WT to PapD, and/or causes a prevention, inhibition or enhancement of the binding of the fusion peptide MBP-G1'-140' to PapD and/or causes a  
25 prevention, inhibition or enhancement of the binding of the peptide G125'-140' (which have the sequence SEQ ID NO: 20) to PapD and/or is capable of inhibiting the restoration of the PapD-PapG complex normally caused by the addition of access PapD.
- 30 The assay used to establish that a substance exhibits one of the above effects is preferably one of the assays described in the examples herein. Of course, also other assays as those discussed above in the methods of the invention may be uti-

lized in order to establish that the compound actually is capable of inhibiting pilus assembly.

One important point which should be taken into consideration when setting up an assay, is the role the chaperones are playing *in vivo*. They bind to the pilus subunits already during the transport through the cell membrane of the subunits and it is therefore assumed, that the pilus subunits are more or less unfolded (i.e. not in their final folded conformation) when they bind to the chaperone. It has been observed by the inventors that the kinetics of binding between completely folded pilus subunits (or analogues thereof) and the chaperone *PapD* is a slow process, although it is known that the process of pilus assembly is relatively fast *in vivo*. In order to speed up the rate of assembly of the chaperone/pilus subunit complex *in vitro* it is contemplated that more or less severe denaturing conditions could be imposed on the pilus subunits (or the analogues thereof) prior to the assay. Such denaturing could be obtained by subjecting the pilus subunits to physical stress (e.g. to elevated temperature, pressure changes, radiation etc.) or to changes in the chemical environment (e.g. changes in ionic strength, changes in pH, or the addition of denaturing compounds or disulphide reducing compounds.

The expression "denaturing compound" refers to a compound which when present as one of the solutes in a liquid phase comprising polypeptide molecules may destabilize folded states of the polypeptide molecules leading to partial or complete unfolding of the polypeptide chains. The denaturing effect exerted by a denaturing compound increases with increasing concentration of the denaturing compound in the solution, but may furthermore be enhanced or moderated due to the presence of other solutes in the solution, or by changes in physical parameters, e.g. temperature or pressure.



As examples of suitable denaturing compounds to be used may be mentioned urea, guanidine-HCl, di-C<sub>1-6</sub>alkylformamides such as dimethylformamide and di-C<sub>1-6</sub>-alkylsulphones.

5 Examples of disulphide reducing compounds are glutathionyl-2-thiopyridyl disulphide, 2-thiocholy-2-thiopyridyl disulphide, 2-mercaptoethanol-2-thiopyridyl disulphide and mercaptoacetate-2-thiopyridyl disulphide.

10 A set of observations confirm the assumption that the pilus subunits are more or less unfolded while bound to the chaperone: As described in Kuehn et al., 1991, it is possible to restore the PapD-PapG complex after denaturation by adding access PapD. Further, preliminary results obtained by capillary electrophoresis show that denaturation of the fusion protein MBP-G1'-140' gives a single form of the fusion protein, whereas two forms of the fusion protein are observed in  
15 capillary electrophoresis before denaturation, one major which is unable to interact with PapD and one minor which is capable of interacting with PapD. It is therefore contemplated that the denatured fusion protein may serve as a superior substrate in the different competitive assays described herein as does the non-denatured form of the fusion protein.

25 Thus, in a preferred embodiment of the invention, the compound of the invention causes a prevention, inhibition or enhancement of the binding of a denatured form of either a pilus subunit or an analogue thereof to PapD, and/or causes a prevention, inhibition or enhancement of the binding of a denatured form of MBP-G1'-140' to PapD and/or causes a prevention, inhibition or enhancement of the binding of a denatured form of G125'-140' to PapD.  
30

As will appear from the examples, compounds which should be capable of interacting with PapD and other chaperones have already been identified and synthesized. These compounds

which are all pyranosides are also an important part of the invention:

- Ethyl 2,3-O-Dibenzoyl-4-O-benzyl-1-thio- $\beta$ -D-glucopyranoside;
- 5 Ethyl 6-O-acetyl-2,3-O-dibenzoyl-4-O-benzyl-1-thio- $\beta$ -D-glucopyranoside;
- Methylglycolyl 6-O-acetyl-2,3-O-dibenzoyl-4-O-benzyl- $\beta$ -D-glucopyranoside;
- 2- (Hydroxy)ethyl 4-O-benzyl- $\beta$ -D-glucopyranoside;
- 10 Sodium glycolyl 4-O-benzyl- $\beta$ -D-glucopyranoside;
- Methyl 2-O-ethyl-4,6-O- (4'-methoxy)phenylmethylene- $\alpha$ -D-mannohexopyranoside;
- Methyl 2-O-ethyl-3-O-dimethyl-t-butylsilyl-4,6-O- (4'-methoxy)phenylmethylene- $\alpha$ -D-mannohexopyranoside;
- 15 Methyl 2-O-ethyl-3-O-dimethyl-t-butylsilyl-4,-O- (4'-methoxy)-benzyl- $\alpha$ -D-mannohexopyranoside;
- methyl 2-O-ethyl-3-O-dimethyl-t-butylsilyl-6-O- (4'-methoxy)-benzyl- $\alpha$ -D-mannohexopyranoside;
- Methyl 2-O-ethyl-3-O-dimethyl-t-butylsilyl-4,-O- (4'-methoxy)-benzyl-6(S)-phenyl- $\alpha$ -D-mannohexopyranoside;
- 20 Methyl 2,3-anhydro-4,6-O-p-methoxybenzylidene- $\alpha$ -D-mannopyranoside;
- Methyl 3-azido-4,6-O-p-methoxybenzylidene- $\alpha$ -D-altropyranoside;
- 25 Methyl 3-azido-2-O-ethyl-4,6-O-p-methoxybenzylidene- $\alpha$ -D-altropyranoside;
- Methyl 3-azido-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl- $\alpha$ -D-altropyranoside;
- Methyl 3-azido-6-O-benzoyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl- $\alpha$ -D-altropyranoside;
- 30 Methyl 6-O-benzoyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl-3-sulfamino- $\alpha$ -D-altropyranoside sodium salt;
- Methyl 6-O-benzoyl-3-deoxy-2-O-ethyl-3-sulfamino- $\alpha$ -D-altropyranoside ammonium salt;
- 35 Methyl 3-azido-6-O-pivaloyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl- $\alpha$ -D-altropyranoside;

- Methyl 6-O-pivaloyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl-3-sulfamino- $\alpha$ -D-altropyranoside sodium salt;  
Methyl 6-O-pivaloyl-3-deoxy-2-O-ethyl-3-sulfamino- $\alpha$ -D-altropyranoside ammonium salt;  
5 Methyl 6-O-pivaloyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl-3-tbutyloxamido- $\alpha$ -D-altropyranoside;  
Methyl 6-O-pivaloyl-3-deoxy-2-O-ethyl-3-oxamido- $\alpha$ -D-altropyranoside ammonium salt;  
Methyl 3-azido-6-O-pyrrol-3'-ylcarboxyl-3-deoxy-2-O-ethyl-4-  
10 O-p-methoxybenzyl- $\alpha$ -D-altropyranoside; and  
Methyl 6-O-pyrrol-3'-ylcarboxyl-3-deoxy-2-O-ethyl-3-sulfamino- $\alpha$ -D-altropyranoside ammonium salt.

Other compounds are of course also possible as interactors with sites in chaperones. As is evident from example 7,  
15 modified peptides also may prove to be useful in the methods of the invention.

As will be clear from the above, the identification of a site in the chaperone which may be affected so as to interfere with pilus assembly is a critical starting point in the  
20 methods described herein for the identification, isolation and synthesis of compounds capable of interacting with periplasmic chaperones.

Thus, the invention also relates to a method for identifying a binding site in a molecular chaperone, comprising

- 25 co-crystallizing the periplasmic molecular chaperone or an analogue thereof with a ligand binding to the periplasmic molecular chaperone or the analogue thereof,

resolving the three-dimensional structure of the chaperone/ligand interaction, thereby resolving the three-dimensional structure of the periplasmic molecular chaperone  
30 or the analogue thereof when binding to the ligand,

determining the site-point(s) involved in the intermolecular interaction between the periplasmic molecular chaperone or the analogue thereof and the ligand, and

- 5 identifying the thus determined site-point(s) of the periplasmic molecular chaperone or the analogue thereof as a binding site in the periplasmic molecular chaperone or the analogue thereof.

By the term "ligand" as used herein, is meant a substance  
10 which exhibit binding to a host or receptor molecule (in this connection a chaperone). The binding is not a non-specific interaction, which means that a binding motif between the ligand and the host or receptor molecule exists. In other words, when bringing a sample of the ligand and a sample of  
15 the host or receptor molecule in contact with each other the complexes formed between the ligand and the host or receptor molecule will substantially all reflect the same intermolecular interactions.

As mentioned above one embodiment of the invention is to  
20 administer a substance which is capable of preventing, inhibiting or enhancing binding between pilus subunits and molecular chaperones.

Accordingly the invention relates to a pharmaceutical composition, comprising, as an active compound, a substance  
25 capable of interacting with at least one type of periplasmic molecular chaperone which binds pilus subunits during transport of these pilus subunits through the periplasmic space and/or during the process of assembly of the intact pilus, in such a manner that binding of pilus subunits to the periplasmic molecular chaperone is prevented, inhibited or enhanced,  
30 in combination with at least one pharmaceutically acceptable carrier or excipient. Preferably such a substance is a substance according to the invention or a substance identified/designed according to the methods of the invention.

The pharmaceuticals and pharmaceuticals discussed herein are thus, according to the invention, for the treatment and/or prophylaxis of the same conditions as those discussed when disclosing the method of treatment of the invention and  
5 caused by the same bacterial species.

Therefore, a pharmaceutical composition, comprising, as an active compound, a substance used in the therapeutic methods of the invention, a substance according to the invention or a substance identified according to the methods of the inven-  
10 tion, in combination with at least one pharmaceutically acceptable carrier or excipient, is a part of the invention.

Such pharmaceutical compositions of the invention could also comprise at least one additional pharmaceutical substance, which i.a. could enhance the pharmaceutical effects exerted  
15 by the substance of the invention or the substance identified according to the invention.

Additional pharmaceutical substances could be steroid hormones, disinfectants, anti-pyretics, etc. Preferably such an additional substance could be an antibacterial agent.

20 Such an anti-bacterial agent could conveniently be selected from the group consisting of penicillins, cephalosporins, aminoglycosides, sulfonamides, tetracyclines, chloramphenicol, polymyxins, antimycobacterial drugs and urinary antiseptics.

25 The invention also relates to a substance employed in the methods of the invention as well as a substance of the invention and a substance identified according to the methods of the invention for use as a pharmaceutical. It is preferred that the pharmaceutical is for anti-bacterial treatment  
30 and/or prophylaxis, especially treatment and/or prophylaxis of diseases caused by tissue-adhering pilus forming bacteria, and it is especially preferred that the pharmaceutical is for treatment and/or prophylaxis of urinary tract infection.

Finally, the invention relates to the use of a substance employed in the methods of the invention as well as a substance of the invention and of a substance identified according to the methods of the invention for the preparation of a pharmaceutical composition for the treatment and/or prophylaxis of bacterial infection.

## LEGENDS TO FIGURES

Fig. 1: *PapD* binds to pilus subunit-related peptides coated in ELISA microtiter plates and in solution. The sequences of the purified, synthetic peptides used in this study are described separately in Flemmer K, Walse B, Roy S, and Kihlberg J, manuscript in preparation. A and B: The peptides from table 1 were coated on microtiter wells and tested for their ability to bind to *PapD* in an ELISA assay described in example 2. Water insoluble peptides were dissolved in 2.5% acetic acid (AA) which had no effect on the binding of G1'-19'WT to *PapD* in this assay. Each graph represents the average of duplicate wells.

Fig. 2: Binding of pilus subunit-related peptides provides protection against enzymatic proteolysis and blocks binding of *PapD* to *PapG*.

A Upper: *PapD* (15  $\mu$ g) was incubated with PBS (lane D) or with 1.5  $\mu$ g trypsin (lane D and Tr) at 37°C for 20 minutes and applied to 20% SDS-PAGE. Coomassie blue stained bands corresponding to full length *PapD* (*PapD*), trypsin (Tr, the NH<sub>2</sub>-terminal fragment of *PapD* (N), and the COOH-terminal fragment of *PapD* starting at residue 100 (C) are indicated).

A Lower: The rate of trypsin cleavage of *PapD* decreased upon preincubation with the G1'-19'WT, G1'-16'WT or K1'-19'WT peptides, but the G2'-21'a-mide peptide had no effect. 50  $\mu$ g of purified *PapD* was preincubated for 15 min at 25°C with a 20-fold

5 molar excess of G1'-19'WT, G1'-16'WT, K1'-19'WT or  
G2'-21'amide peptides (any trypsin cleavage sites  
in these peptides occur at their amino termini  
where they were predicted not to interfere with  
PapD), or an equivalent volume of water. Each  
sample was then incubated at 37°C with 3.2 µg of  
trypsin. Aliquots were removed after 0, 5, 10, 20,  
30 and 40 min and boiled in SDS-PAGE sample buffer  
to stop the digestion. The samples were applied to  
10 15% SDS-PAGE gels and the amount of full-length  
PapD remaining was determined by densitometry of  
Coomassie blue stained bands corresponding to full-  
length PapD which are shown.

15 B: PapD incubated with the G1'-19'WT, G1'-16'WT,  
K1'-19'WT, and G2'-21'amide was added to reduced,  
denatured PapD-PapG and the amount of PapD-PapG  
complex restored in each sample was quantitated as  
described in example 2. The % inhibition represents  
the amount of the PapD-PapG complex restored with  
20 peptide-treated PapD compared to the amount of  
PapD-PapG complex restored with untreated PapD. The  
graph represent the average of 4 experiments. PapD  
and PapD-PapG complex were prepared as described in  
F. Lindberg et al, J. Bacteriol. 171, 6052 (1989)  
25 and S.J. Hultgren et al., Proc. Natl. Acad. Sci.  
USA 86, 4357 (1989).

Fig. 3 Stereoscopic view of the three-dimensional struc-  
ture of PapD co-crystallized with the G1-19WT  
peptide determined to 3.0 Å resolution. The peptide  
30 is bound in an extended conformation along the G1 β  
strand in the cleft of PapD with the terminal car-  
boxylate group forming hydrogen bonds with residues  
Arg-8 and Lys-112 of PapD.

Fig. 4 Stereoscopic view of the 3.0 Å resolution electron  
density of the G1-19WT peptide COOH-terminus and  
35 neighbouring PapD residues, superimposed on the

refined structure. The electron density map was calculated with coefficients  $(2|F_o| - |F_c|)$  and contoured at  $1\sigma$ .

- Fig. 5 Stereoscopic view of a PapD-peptide complex showing its interaction with a twofold non-crystallographically related complex.
- Fig. 6 Superposition of COOH-terminal domains of native PapD (thick lines) and complexed PapD (thin lines) showing the  $13^\circ$  difference in the hinge bend angle between the two structures. The structures were superimposed using the LSQ option the program O. The resulting rms for 98 Ca atoms of the COOH terminal domain was 0.66 Å.
- Fig. 7: Ability of PapD Arg-8, Lys-112 and Thr-7 mutants to bind PapG and restore the PapD-PapG complex in vitro. 0.4 µg of PapD-PapG complex was reduced and denatured as described in Fig. 2B. The denatured PapD-PapG complex was then diluted with 0-630 ng purified wild type (WT) or mutant PapD. The amount of PapD-PapG restored in the samples was determined as in Fig. 2B and was graphed as a percentage of the greatest amount of PapD-PapG restored. The values graphed represent the average of 3 experiments. Wild type and mutant PapD and the PapD-PapG complex were purified as described in F. Lindberg et al, J. Bacteriol. 171, 6052 (1989) and S.J. Hultgren et al., Proc. Natl. Acad. Sci. USA 86, 4357 (1989).
- Fig. 8: 6-hydroxydopamine, hdo\_0, and other members of the hdo family. hdo\_0 Was used as starting-point for the construction of a family of compounds intended to bind to the binding-site of PapD. hdo\_4: "Aryl" denotes 3,4-dihydroxyphenyl. hdo\_6: "Aryl" is a 3,4-dihydroxyphenyl group.



hdo\_7: "Aryl" is a 3-hydroxyphenyl group.

hdo\_8: "Aryl" is a 3-nitrophenyl group.

hdo\_9: "Aryl" is a phenyl group.

Fig. 9: The hdo family and its interaction with PapD.

5 A: "Aryl" is a phenyl group, a heteroaromatic ring, or a substituted phenyl group with polar substituents unsymmetrically in order to obtain hydrophobic contact with the protein on one side and interaction with solvent water on the other.

10 B: Examples of "Aryl" wherein R is hydrogen or alkyl; "X" is an oxygen atom, a sulphur atom, or an amino group; and "CG" is a negatively charged group, such as a carboxyl group, a tetrazoyl group, a phosphate group, a phosphate ester or a sulphate group.

15 C: Examples of the group CG.

Fig. 10: General reaction scheme for the production of various members of the hdo family.

20 Fig. 11: General reaction scheme for the production of various members of the bpy family.

Fig. 12: General reaction scheme for the production of various members of the bpy family.

Fig. 13: General reaction scheme for the production of various members of the bpy family.

25 Fig. 14: General reaction scheme for the production of various members of the bpy family.

Fig. 15: General reaction scheme for the production of various members of the bpy family.

30 Fig. 16: General reaction scheme for the production of various members of the bpy family.

Fig. 17: Graph over the inhibitory effect of different lengths of C-terminal peptide fragments of *PapG* on the binding between *PapD* and the fusion protein MBP-G1'-140'.

5 The G1'-8' fragment exhibits a significantly higher inhibitive effect on the binding than does the shorter G1'-6' and G1'-7' fragments.

Fig. 18: Graph over the inhibitory effect of different analogues of G1'-8' (wherein the 8 amino acid residues one at a time has been replaced by serine (S) or Alanine (A)) on the binding between *PapD* and the fusion protein MBP-G1'-140'.

10

15 The effect of the replacements of residues 4', 5' and 6' in G1'-8' reveals that these residues are important in the interaction with *PapD*, as the replacement of these residues results in less inhibitory peptides.

Fig. 19: Graph over the inhibitory effect of different analogues of G1'-8' (wherein the 8 amino acid residues one at a time has been deleted simultaneously with the addition of an N-terminal serine) on the binding between *PapD* and the fusion protein MBP-G1'-140'.

20

25 Also in this experiment, the significance of amino acid residues 4', 5', and 6' are emphasized, as the deletions at these positions leads to less inhibitory peptides.

Fig. 20: MBP/G fusion constructs, *PapG*-truncates and synthetic peptides used in example 10.

30

35

The open box indicates the primary sequence of *PapG*. The positions of the four Cys residues are shown. The hatched bar represents MBP. The starting and ending residues of *PapG* fused onto the COOH-terminus of MBP are noted for each fusion. The terminating residue for each *PapG*-truncate is also

indicated. The names of the MBP/G fusion proteins are listed. In the lower portion of the figure, the solid boxes localize the *PapD* interactive sites and the sequences of the four peptides used in the example are listed (cf. also SEQ ID NOS: 19-22).

Fig. 21: *PapD*-MBP/G interactions in vivo.

Periplasmic extracts containing *PapD* and each MBP/G fusion were subjected to amylose affinity chromatography. The eluates were analyzed on A) a 12.5% Coomassie blue stained SDS polyacrylamide gel; B) a western blot using anti-*PapD* antiserum; or C) a silver stained IEF gel. In Figs. 21A, B and C, samples were purified from periplasmic extracts containing *PapD* and MBP (lane 1), *PapD* and MBP-G1'-19' (lane 2), *PapD* and MBP-G1'-81' (lane 3), *PapD* and MBP-G1'-140' (lane 4). The position of co-purified *PapD* is indicated. MBP alone and MBP/G fusion truncates also co-purified with the MBP/G fusions. The molecular weight of the slowest migrating bands on SDS-PAGE correspond to each full length MBP/G fusion protein. On the IEF gel (C) several bands could be seen for the same reasons. A unique band at pI 5.2 was detected in Fig. 2C, lane 4. This band was excised, boiled in SDS sample buffer and analyzed by western blotting with anti-*PapD* and anti-MBP antisera. It was composed of both MBP-G1'-140' and *PapD* (Fig. 2D, lane 1).

Fig. 22: Inhibition of chaperone function by expression of MBP/G fusions in vivo.

Strains carrying pFJ22 (*papDJKEFGA*) and plasmids encoding the MBP/G fusions were induced with 1 mM IPTG. Periplasmic extracts containing pilus subunits and MBP (lane 1), or MBP-G1'-19' (lane 2), or MBP-G1'-81' (lane 3), or MBP-G1'-140' (lane 4) were analyzed by western blotting using anti-*PapA* antiserum (A), or anti-*PapD*-*PapG* antiserum (B), or

anti-tip fibrillum antiserum (C). In this assay the presence of the subunit indicates that chaperone-subunit interactions occurred since subunits are degraded in the absence of an interaction with the chaperone. Note that the presence of the *PapA*, *PapG* and *PapF* subunits decrease significantly when co-expressed with MBP-G1'-140'.

Fig. 23: Binding of *PapD* to MBP/G fusion proteins in vitro. (A) *PapD* was incubated with 1  $\mu$ g of amylose affinity purified MBP (lane 1), MBP-G1'-19' (lane 2), MBP-G1'-81' (lane 3), or MBP-G1'-140' (lane 4) and complex formation was evaluated on the silver stained IEF gel. Positions of MBP/G fusions, *PapD* and the *PapD*-MBP-G1'-140' complex are indicated. (B) Amylose affinity purified MBP/G fusion proteins were coated to wells of microtiter plates. The concentration of the MBP/G proteins in the wells is indicated. Binding of 50 pmol/50 ml of *PapD* to the immobilized proteins was determined by ELISA using anti-*PapD* antiserum.

Fig. 24: Identification of *PapD*-*PapG2* truncate complex by acidic native gel electrophoresis. Periplasm containing *PapD* and *PapG*, *PapG2* or *PapG3* (lane 2, 3 and 4 respectively) were subjected to Gal  $\alpha$ (1-4) Gal chromatography and eth eluates were analyzed on acidic native gel electrophoresis following western blotting using anti-*PapD* (A) and anti-*PapG* (B) antisera. Purified *PapD*-*PapG* complex was loaded in lane 1.

Fig. 25: Characterization of a second site on *PapG* recognized by *PapD*. Four synthetic peptides overlapping the *PapG* region from residue 156' to 120' (indicated in Fig. 20) were coated to wells of microtiter plates. The concentration of each peptide in the wells is indi-

cated. Binding of 200 pmol/ 50 ml of *PapD* to the immobilized peptide was determined by ELISA using anti-*PapD* antiserum.

## EXAMPLES

### 5 EXAMPLE 1

*Identification of the motif of binding between PapD and G1'-19'WT*

#### *Materials and Methods*

*PapD* was prepared as described by A. Holmgren et al. in J. Mol. Biol. 203, 279 (1988) and obtained from the Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, USA.

The peptide G1'-19'WT was prepared using Fmoc solid phase synthesis followed by purification by reversed phase HPLC. The peptide was obtained from Department of Chemistry, University of Lund, Lund, Sweden. Table 1 indicates the peptides used in binding assays in this application.

In the present application, amino acid residues or peptide fragments originating from other peptides than *PapD* (e.g. from *PapG* or from *PapK*) will be indicated with ', e.g. "G1'-19'WT" or "Pro-1'", in order to distinguish such residues from amino acid residues of *PapD*. Further, the numbering is from the C-terminal end of the non-*PapD* peptide, i.e. G1'-19' denotes the 19 C-terminal amino acid residues of *PapG*.

### 25 Crystallisation of *PapD*-peptide complex

Crystals of the *PapD*-peptide complex were obtained by vapour diffusion against 20% PEG8000, 0.1 M cacodylate buffer at pH 5.0, and 0.2 M calcium acetate. The crystallisation drop contained equal volumes of reservoir and protein solutions. The protein solution (17 mg/ml) contained a 1:1 molar ratio

of PapD and peptide in 20 mM MES (2-[N-Morpholino]ethane sulphonic acid) at pH 6.5 with 1.0%  $\beta$ -octyl glucoside.

### *X-ray crystallography*

The obtained crystals of the complex between PapD and the protein were mounted inside sealed quartz-glass capillary tubes and initially characterised by examining them on a X-ray precession camera thereby obtaining pictures of the X-ray diffraction pattern on photographic film representing spatially non-distorted images of the reciprocal lattice of the crystals. Using standard analysis of the images, it was determined that the crystals have a monoclinic space group, C2, with cell dimensions:  $a = 130.7 \text{ \AA}$ ,  $b = 83.5 \text{ \AA}$ ,  $c = 59.2 \text{ \AA}$ , and  $\beta = 117.2^\circ$ , two molecules in the asymmetric unit and diffract to  $2.9 \text{ \AA}$  resolution on a lab X-ray source with a rotating anode and Cu  $K_\alpha$  target.

### *Collection and processing of experimental data*

In order to solve the atomic structure of molecules by X-ray crystallography, the positions and intensities of the diffraction maxima have to be measured. Intensity data for the PapD-peptide crystals were collected on a Huang-Hamlin multi-wire area-detector system (Xuong et al., 1985). All data were obtained from a single crystal and initially processing was carried out using the MADNES software package (Messerschmit and Pflugrath, 1987). Merging and scaling of the data was carried out using ROTAVATA and AGROVATA from the CCP4 package (see CCP4, The SERC (UK) Collaborative Project No. 4, A Suite of Programs for Protein Crystallography, Daresbury Laboratory, UK). The final data set contained 9592 independent reflections with an  $R_{\text{sym}}$  of 6.8% for data between 20.0 and  $3.0 \text{ \AA}$  resolution ( $R_{\text{sym}} = \frac{\sum \sum (|I_{hi}| - |I_h|)}{\sum_{hi} I_{hi}}$ , where  $I_{hi}$  and  $I_h$  are the

intensities of the individual and mean structure factors, respectively).

### *Solution of three-dimensional structure*

The structure of the PapD-peptide complex was solved using the program XPLOR performing the standard method of molecular replacement. The search model used was the refined 2.0 Å resolution structure of PapD. Using 8.0 Å to 4.0 Å resolution data, the self-rotation function gave a clear non-crystallographic two-fold axis, and the top two peaks of the cross-rotation gave the correct solution which was improved using Patterson correlation (PC) refinement. The correct solution was also obtained from the top peaks in the translation functions, the R-factor being 39.0% for 8.0 to 4.0 Å resolution data (the R-factor defined as  $R\text{-factor} = \frac{\sum |F_o| - |F_c|}{\sum F_o}$ ).

Refinement was obtained with subsequent rigid body refinement in which all four domains of the 2 PapD molecules in the asymmetric unit were allowed to refine independently resulting in a R-factor of 36.4% for the same data.

Examination of an  $|F_o| - |F_c|$  electron density map at this stage using the graphics program O showed clear density corresponding to the peptide in the PapD cleft and running along the surface of the protein. The orientation of the peptide was easily determined from the electron density, but initially only the final 10 C-terminal amino-acids of the peptide could be modelled into density.

### *Refinement and analysis of structure.*

Simulated annealing refinement with XPLOR was initiated at this stage. Several additional cycles of model building and refinement were carried out with a further 4 peptide amino-acids being added to yield an R-factor for the current model of 18.2% for 8.0 to 3.0 Å resolution data. The model at the present stage of refinement (which contains no water molecules and does not include the first 5 N-terminal amino-acids of the peptide) has root-mean square (rms) deviations from

ideal geometry of 0.020 for bonds lengths and 4.2° for bond angles.

The peptide is seen to bind in an extended conformation with the C-terminal pro-1' anchored within the inter-domain cleft and proposed (6) subunit binding site. Hydrogen bonds are formed between the peptide carboxy terminus and two invariant positively charged residues of *PapD*, Arg-8 and Lys-112. Site-directed mutagenesis has now confirmed that Arg-8 and Lys-112 are essential for the binding of pilin subunits both *in vitro* and *in vivo* (cf. example 2)). The Pro-1' side chain makes van der Waals contacts in the cleft with residues from both domains of *PapD*: Thr-7, Thr-152, Ile-154, Thr-170 and Ile-194. The neighbouring peptide residue, Phe-2', lies in a shallow pocket formed between the two  $\beta$ -sheets of the N-terminal domain and makes hydrophobic interactions with Leu-4, Thr-7, Thr-109 and Ile-111. The peptide then runs along the surface of the N-terminal domain, forming a parallel  $\beta$ -strand interaction with strand G1. In this way 7 main chain hydrogen bonds are formed between Met-8' to Phe-2' of the peptide and Gln-104 to Lys-110 of *PapD*, thus extending the  $\beta$ -sheet of *PapD* out into the peptide.

Apart from the C-terminal residues Phe-2' and Pro-1' there are relatively few contacts between the side chains of the peptide and *PapD*. The major interactions are provided by the main chain hydrogen bonds to strand G1. There are, however, a number of hydrophobic interactions within the  $\beta$ -sheet, in particular between the peptide's Met-8' and Met-6' with Leu-103, Ile-105 and Leu-107 of strand G1.

Calculations revealed that the four hydrophobic peptide side chains of residues 2', 4', 6', and 8' contribute 20% of the total buried surface area (582 Å<sup>2</sup>) between the peptide and the protein. Therefore, even though the major stabilization of the complex is provided via hydrogen bonding, hydrophobic interactions are not insignificant, and it is believed that they provide part of the explanation for the specificity of



- PapD* for pilus related peptides and subunits. Experimental support of this theory is provided by the reduced binding of *PapD* to the peptide G2'-21'amide as compared to the G1'-19'WT peptide (Fig. 1B and table C). Hydrogen bonding of the COOH-terminus of the G2'-21'amide (which lacks Pro-1') to Arg-8 and Lys-112 of *PapD* allows main chain hydrogen bonding, but dislocates the four hydrophobic side chains in the peptide from their subsites in *PapD*, resulting in a reduction in binding strength.
- 10 Within the crystal the *PapD*-peptide  $\beta$  sheet was extended even further as a result of non-crystallographic twofold symmetry which placed a second *PapD*-peptide complex adjacent to the first so that the two bound peptide chains interacted as antiparallel  $\beta$  strands. In the present model, eight main
- 15 chain hydrogen bonds are formed between the two peptides; see the following table:

Table B

<i>PapD</i> <sub>1</sub>	Peptide <sub>1</sub>	Peptide <sub>2</sub>	<i>PapD</i> <sub>2</sub>
		Pro-1'-COOH	Arg-8-NH <sub>2</sub>
		Pro-1'-COOH	Lys-112-NZ
		Phe-2'-NH	Lys-110-C=O
	Glu-13'-NH	Ser-3'-C=O	
	Glu-13'-C=O	Ser-3'-NH	
		Leu-4'-C=O	Lys-110-NH
		Leu-4'-NH	Gln-108-C=O
	Ser-11'-NH	Val-5'-C=O	
	Ser-11'-C=O	Val-5'-NH	
		Met-6'-C=O	Gln-108-NH
		Met-6'-NH	Ala-106-C=O
	Ser-9'-NH	?	Thr-7'-C=O
	Ser-9'-C=O	?	Thr-7'-NH
Gln-104-C=O	Met-8'-NH	Met-8'-C=O	Ala-106-NH
Ala-106-NH	Met-8'-C=O	Met-8'-NH	Gln-104-C=O
	Thr-7'-NH	?	Ser-9'-C=O
	Thr-7'-C=O	?	Ser-9'-NH
5	Ala-106-C=O	Met-6'-NH	
	Gln-108-NH	Met-6'-C=O	
		Val-5'-NH	Ser-11'-C=O
		Val-5'-C=O	Ser-11'-NH
	Gln-108-C=O	Leu-4'-NH	
	Lys-110-NH	Leu-4'-C=O	
		Ser-3'-NH	Glu-13'-C=O
		Ser-3'-C=O	Glu-13'-NH
	Lys-110-C=O	Phe-2'-NH	
10	Lys-112-NZ	Pro-1'-COOH	
	Arg-8-NH <sub>2</sub>	Pro-1'-COOH	

A mixed  $\beta$  sheet is thus created between the two complexes and extends over ten  $\beta$  strands (Fig. 5). No contacts are observed between the two non-crystallographically related *PapD* molecules themselves, both of which were positioned in similar environments within the crystal and possessed a similar

number of intermolecular contacts. The calculated buried surface area between the two non-crystallographically related peptides was 520 Å<sup>2</sup>, a value similar to the surface area buried between the protein and peptide. This "dimerization" appears to be a consequence of crystal packing, because all evidence shows that PapD forms monomeric complexes with peptides or with intact PapG in solution.

The hydrogen bonding pattern between PapD and the peptide breaks at Ser-9', but the peptide remains in van der Waals contact with PapD until Ser-11' where the peptide runs beyond the F1-G1 loop, but remains hydrogen bonded to the non-crystallographically related peptide as far as Glu-13' (table B). The last resolved amino acid of the peptide is Gly-14' which is positioned close to the binding cleft of the non-crystallographically related PapD. the first five amino terminal amino acids, including three positively charged residues, had no density and therefore must have been disordered in the crystal structure. In agreement with this, the peptide's amino terminus is not important for binding to PapD in solution, since the G1'-7'WT peptide lacking the amino terminal 12 amino acids is found to be an effective inhibitor of PapD binding to the immobilized G1'-19'WT peptide (table C and example 2).

The structures of the individual PapD domains in the peptide complex are essentially the same as those of native PapD (Holmgren and Bränden, 1989). However, there is a significant movement of the domains with respect to each other with a 13° jaw-closing or hinge bending motion making the angle of the PapD boomerang more acute (Fig. 6). Whether or not this conformational change is the result of binding of peptide or different crystal packing between the two crystal structures is unclear.

In the native PapD structure the electron density obtained for the long F1-G1 loop is poor between residues 96 and 102, suggesting that it is rather flexible and disordered in the

crystal. In the peptide complex, however, this loop is better resolved, indication that binding of the peptide makes this loop more rigid. Superimposing the NH<sub>2</sub>-terminal domains of native *PapD* and the peptide complex shows that there is also a significant difference in the F1-G1 loop position between the two structures (rms for the 110 NH<sub>2</sub>-terminal C $\alpha$  atoms is 1.84 Å, with a maximum main chain movement of about 9 Å for Leu-103). In the peptide complex the loop is seen to twist at one end away from the  $\beta$  barrel of the NH<sub>2</sub>-terminal domain thus facilitating a more extensive contact between strand G1 and the peptide. As with the hinge-bending of the two domains, it is not yet possible to say with certainty whether this loop shift is a consequence of peptide binding or of crystal packing; the rather open conformation of the F1-G1 loop suggests that it may be largely the latter. Nevertheless, evidence that similar interaction between *PapD* and peptides or pilus subunits occur in solution is provided by protease protection experiments after the F1-G1 loop of *PapD* is protected from tryptic cleavage by the binding of both native *PapG* and the G1'-19'WT peptide (example 2 and Fig. 2A).

#### EXAMPLE 2

*Binding between PapD and the carboxyl terminus of other Pap peptides*

G1'-19'WT, E1'-19'WT, F1'-19'WT, K1'-19'WT and H1'-19'WT peptides (see table 1) were synthesized corresponding to the 19 carboxyl terminal residues of P pilus subunit proteins *PapG*, *PapE*, *PapF*, *PapK* and *PapH*, respectively (see: Grant et al., 1992, and references cited therein).

Table 1

	Name	SEQ ID NO:	Peptide sequence
	G1'-19'WT	6	NH <sub>2</sub> -GKRKPGELSGSMTMVLSFP-COOH
	G1'-16'WT	7	NH <sub>2</sub> -KPGELSGSMTMVLSFP-COOH
5	G1'-11'WT*	8	NH <sub>2</sub> -SGSMTMVLSFP-COOH
	G1'-7'WT	9	NH <sub>2</sub> -TMVLSFP-COOH
	G1'-19'SV	10	NH <sub>2</sub> -GKRKPVELSGSMTMVLS <del>SP</del> -COOH
	G1'-19'amide	11	NH <sub>2</sub> -GKRKPGELSGSMTMVLSFP- <u>CONH<sub>2</sub></u>
	G2'-21'amide	12	NH <sub>2</sub> - <del>EE</del> GKRKPGELSGSMTMVLSF- <u>CONH<sub>2</sub></u>
10	E1'-19'WT*	13	NH <sub>2</sub> -QNLIAGPFSATATLVASYS-COOH
	H1'-19'WT	14	NH <sub>2</sub> -KKLEAGNYFAVLGFRVDYE-COOH
	K1'-19'WT	15	NH <sub>2</sub> -KSVVPGDYEATATFELTYR-COOH
	F1'-19'WT	16	NH <sub>2</sub> -GILNGGDFQTTASMAMIYN-COOH
	MS	17	NH <sub>2</sub> -YALAPNAV IPTSLALL-COOH

15 \* Water insoluble peptides

The residues in the peptides were numbered starting with the carboxyl terminal residue as 1, and ending with the amino terminal residue. Peptides were also synthesized which deviated in length (G1'-16'WT, G1'-11'WT, G1'-7'WT) or sequence (G2'-21'amide, G1'-19'SV, G1'-19'NH<sub>2</sub>) from the wild type PapG carboxyl terminal sequence. Using an enzyme-linked immunosorbent assay (ELISA), we tested the ability of PapD to bind to each peptide coated on wells of microtiter plates:

5 mg/ml stock solutions of peptides in water or 50% acetic acid were diluted to a concentration of 2.5 pmol/50  $\mu$ l in PBS. 50  $\mu$ l of the peptide solution was coated overnight onto microtiter wells (Nunc-Immuno Plate Maxisorp) at 4°C. The solutions in the plates were discarded, and the wells were blocked with 200  $\mu$ l 3% bovine serum albumin (BSA, Sigma) in PBS (120 mM NaCl/2.7 mM KCl/10 mM phosphate buffer salts, pH 7.4) for 2 hours at 25°C. The plates were washed vigorously 3

times with PBS and incubated with 50  $\mu$ l of the indicated amount of purified *PapD* (Lindberg et al, 1989). After 3 washes with PBS, the wells were incubated with a 1:500 dilution of anti-*PapD* rabbit antiserum (Lindberg et al, 1989) in 3% BSA/PBS for 45 min. at 25°C. After 3 washes with PBS, the wells were incubated with a 1:1000 dilution of goat anti-rabbit IgG coupled to alkaline phosphatase (Sigma) in 3% BSA/PBS for 45 min. at 25°C. Following 3 washes with PBS and 3 washes with developing buffer (10 mM diethanolamine pH 9.5, 0,5 mM  $MgCl_2$ ), 50  $\mu$ l filtered p-nitrophenyl phosphate substrate (Sigma) in developing buffer (10 mg/ml) was added. The absorbance at 405 nm was read after 60 min of incubation in the dark at 25°C.

In addition, the ability of water-soluble peptides to inhibit *PapD* binding to G1'-19'WT-coated wells in a soluble inhibition ELISA were tested, since peptide conformations may have been affected by binding to the plastic microtiter plates:

Microtiter wells were coated overnight at 4°C with 50  $\mu$ l of 2,5 pmol/50  $\mu$ l of the G1'-19'WT peptide. The wells were washed with PBS and blocked with 3% bovine serum albumin (BSA). A 25-fold molar excess of each test peptide was pre-incubated with 100 pmoles *PapD* for 30 min. and the *PapD*-peptide solution was then added to the coated wells and incubated at 25°C for 45 min in the presence of 3% BSA/PBS. The subsequent primary antibody, secondary antibody and developing steps are described above. The ability of the peptides to inhibit binding of *PapD* to the G1'-19'WT peptide was calculated by dividing the amount of *PapD* binding in the presence of peptide with the amount of *PapD* binding in the presence of water. 0% inhibition includes values where binding was greater than that of *PapD* pre-incubated with  $H_2O$ .

The results of *PapD* binding to wild type, variant and different length peptides are shown in Figs. 1A, B and C.

- As can be seen the peptides bound well to the *PapG* peptide, moderately to *PapE*, *PapF* and *PapK*, whereas there was no binding to *PapH* and the random hydrophobic peptide, MS. (Fig. 1A). These results suggest that the chaperone recognizes
- 5 *PapG*, *PapE*, *PapF* and *PapK* in part by binding to the carboxyl terminus of these subunits. The inability of *PapD* to interact with the *PapH* peptide indicates that *PapD* binds differently to *PapH*, possibly due to *PapH*'s function as a polymerization terminator (Baga et al., 1987). The ability of a peptide to
- 10 inhibit *PapD* binding to immobilized G1'-19'WT generally corresponded to the affinity that *PapD* had for the respective immobilized peptide (Fig. 1, % inhibition), arguing that the interaction of *PapD* with soluble and immobilized peptides is similar.
- 15 As can be seen from example 1, the molecular basis of the *PapD*-peptide interactions have been studied by co-crystallization of *PapD* with the G1'-19'WT peptide. In summary, the 19 amino acid peptide is anchored in the chaperone cleft by hydrogen bonds between the peptide's carboxy terminus and
- 20 Arg-8 and Lys-112 which are invariant in all periplasmic chaperones (Holmgren et al., 1992). The peptide bound to the G1 strand of *PapD* as a parallel  $\beta$ -strand, forming at least 10 backbone hydrogen bonds, and resulted in an ordering of the F1 to G1 loop.
- 25 Replacement of the carboxyl terminal proline on the *PapG* peptide with an amide (G2'-21'amide) abolished binding to *PapD* in solution and reduced binding of *PapD* to the immobilized peptide by approximately 75% (Fig. 1B). In contrast, substituting only the carboxylate group with an amide to
- 30 create the G1'-19'NH<sub>2</sub> peptide did not affect binding to *PapD* in either the immobilized or soluble inhibition ELISA assays (Fig. 1B). These results indicate that the terminal proline probably is required to position the carboxylate group so that it can form hydrogen bonds with the invariant Arg-8 and
- 35 Lys-112 cleft residues.

PapD also binds to the immobilized shorter peptides G1'-16'WT and G1'-11'WT and the G1'-16'WT and G1'-7'WT peptides inhibited binding of PapD to the immobilized G1'-19'WT peptide in solution. Table C shows the ability of a 25-fold molar excess of the water soluble peptides to inhibit binding of 100 pmol/well of PapD to G1'-19'WT-coated wells. The % inhibition represents the percentage of PapD binding in the presence of the peptide compared to in the presence of water and are the average of two experiments performed in duplicate:

Table C

Peptide	% Inhibition
G1'-19'	63
G1'-7'	49
15 G1'-4'	0
MS	1
G2'-21'amide	0
G1'-19'amide	56

PapD does not bind to immobilized G1'-7'WT peptide, probably because this peptide is too short to bind to the microtiter well as well as PapD. It, thus seems that as few as 7 carboxyl terminal residues are necessary for PapD binding of a peptide. Together these results support a model where, in addition to the anchoring interaction of the carboxylate group in the PapD cleft, seven carboxyl terminal residues are required for "zippering" the peptide along the G1 strand of PapD during PapD-peptide complex formation.

Conserved residues phenylalanine and glycine residues at positions 2 and 14 from the carboxyl terminus were substituted by serine and valine, respectively, to create the



peptide G1'-19'SV and decreased binding to *PapD* by 36%. This could result from less efficient coating or presentation in the microtiter plate since G1'-19'SV was as efficient a soluble *PapD* inhibitor as the peptide G1'-19'WT (Fig. 1B).  
5 Since these residues are critical for incorporation of *PapG* into the pilus, it is believed that they be more important for intersubunit polymerization interactions and pilus assembly than in *PapD* subunit interactions *in vivo*.

Partial digestion with trypsin cleave *PapD* in the F1-G1 loop  
10 at residues Leu-103 and Lys-99, respectively:

400  $\mu$ g of *PapD* was partially digested by incubation with either 4.5  $\mu$ g trypsin or 0.45  $\mu$ g chymotrypsin on PBS for 20 min. at 37°C. The *PapD* digests were applied to a C-18 HPLC column (Beckman) and 2 major fragments were eluted with a 0-  
15 100% acetonitrile gradient in 0.01% trifluoroacetic acid. The *PapD* fragments were identified by their molecular weight on SDS-PAGE and amino-terminal sequencing. The N-terminal amino acid sequences of the approximately 14 kDa tryptic and chymotryptic fragments were identified as residues 100-108 and  
20 104-109 of *PapD*, respectively, corresponding to cleavage after Lys-99 (for trypsin) and Leu-103 (for chymotrypsin). The N-terminal sequences of the 11 and 12 kDa bands in the digests were identical to the N-terminal sequence of *PapD*.

The G1'-19'WT, G1'-16'WT and K1'-19'WT peptides, but not the  
25 G2'-21'amide peptide, reduced the rate of tryptic cleavage of *PapD* over time (Fig. 2A). These data argue that binding of *PapD* to the peptide altered contacting the loop and thereby protectin to form cleavage. This effect may be related to the ordering of the F1-G1 loop of *PapD* observed in the *PapD*-  
30 peptide crystal structure, and suggests that, in solution, both peptides extends along the G1  $\beta$ -strand and interact with the F1-G1 loop of *PapD*.

As described in Kuehn et al., Proc. Natl. Acad. Sci. USA 88, 10586 (1991), native *PapD* is able to bind to reduced, denatured *PapG* and restore the *PapD-PapG* complex *in vitro*.

This reconstitution assay was used to determine the ability of the peptides to inhibit *PapD* activity *in vitro*. The limited solubility of the G1'-7'WT peptide unfortunately has prevented the testing of this peptide in the assay. Increasing amounts of the G1'-19'WT, G1'-16'WT and K1'-19'WT peptides inhibit restoration of the *PapD-PapG* complex by *PapD*, while the G2'-21'amide peptide has no effect (Fig. 2B). The ability of the G1'-19'WT, G1'-16'WT and K1'-19'WT peptides to prevent *PapD* from binding to *PapG* indicated that these peptides bound in the subunit binding site of *PapD*.

The ability of peptides to inhibit restoration of the *PapD-PapG* complex by *PapD* was tested as follows.: 0.3  $\mu$ g of *PapD-PapG* complex was reduced and denatured by incubation at 25°C for 20 min with 4 M urea/10 mM dithiothreitol (DTT), 1.2  $\mu$ g (50 pmol) of *PapD* was incubated at 25°C for 10 min with 5-14.5  $\mu$ g (2.5-7.25 nmol) of peptide. The *PapD*-peptide solution was then added to the reduced, denatured *PapD-PapG*, incubated at 25°C for 10 min, and applied to an IEF 3-9 gel (Pharmacia). The amount of *PapD-PapG* restored in each sample was quantitated by densitometry of the silver stained IEF band corresponding to the pI of the DG complex.

Site directed mutations in strictly conserved cleft residues of *PapD* predicted to be critical in the *PapD*-peptide interaction were constructed to test whether the *PapD*-peptide crystal structure is a reflection of part of the *PapD*-pilus subunit interaction interface (the positions of the mutations are indicated in Fig. 3). Highly conserved Thr-7 was changed to a valine (Thr-7-Val) in order to test whether its hydroxyl group formed hydrogen bonds critical for *PapD* binding to subunits.

This mutation removed the hydroxyl group while maintaining the steric volume of the side chain. Mutations in Lys-112 and

Arg-8 were designed to test whether hydrogen bonding to the terminal carboxylate group of pilus subunits is a critical feature of the chaperone recognition process. The invariant Lys-112 residue was changed to an alanine (Lys-112-Ala) to  
5 remove the charged side chain and to a methionine (Lys-112-Met) to replace the charged group with a hydrophilic group while maintaining the side chain packing. The invariant Arg-8 has been shown previously to be required for the ability of *PapD* to bind subunits and mediate pilus assembly *in vivo*  
10 (Slonim et al, 1992). Glu-167 /E167, a variable residue in domain 2 of *PapD* (Fig. 3), does not appear to be involved in the interaction between *PapD* and the peptide in the crystal structure and mutations in this residue have been shown to have little or no effect on *PapD* function *in vivo* (Slonim et  
15 al, 1992). E167 was changed to histidine (E167H) to test whether this negatively charged residue at the lip of domain 2 has any role in the *PapD*-peptide interaction. All of the mutants were secreted into the periplasmic space as stable proteins similar to wild type *PapD*. In addition, the elution  
20 profile from a cation-exchange FPLC column and the electrophoretic properties of the purified mutant *PapDs* were similar to the wild type protein, supporting the prediction that these mutations would not affect the overall structure of *PapD* (data not shown).

25 The ability of the mutant chaperones to bind pilus subunits and modulate pilus assembly *in vivo* was correlated to their ability to bind to the G1'-19'WT peptide and to *PapG* *in vitro*. Wild type *PapD* bound the G1'-19'WT peptide to cause a mobility shift towards the negative electrode in a native  
30 polyacrylamide gel assay (Lam and Calderwood, 1992), probably due to a net positive charge increase in the *PapD*-peptide complex. In contrast, when Arg-8-Gly, Arg-8-Ala and Lys-112-Ala *PapD* mutants were incubated with the G1'-19'WT peptide they did not cause a mobility shift in the native-PAGE assay,  
35 indicating that these mutations in *PapD* abolished peptide binding. Similarly, mutations in Arg-8 and Lys-112 abolished the ability of *PapD* to bind to *PapG* and reconstitute the

*PapD*-*PapG* complex in vitro (Fig. 7) and to bind to pilus subunits and mediate pilus assembly in vivo (cf. table 2).

Table 2

5	Residue		HA	Pilus	Subunit stabilization <sup>#</sup>				
	class	Mutation	titer*	assembly <sup>+</sup>	<i>PapA</i>	<i>PapE</i>	<i>PapG</i>	<i>PapF</i>	<i>PapK</i>
	Invariant	Arg-8-Gly	None	-	-	-	-	-	-
		Arg-8-Ala	None	-	+	-	+	-	+
		Arg-8-Met	None	-	++	-	++	-	++
		Lys-112-Ala	None	-	+	-	+	-	+
		Lys-112-Met	None	-	++	-	-	-	-
	Conserved	Thr-7-Val	64	+++	++++	++++	++++	++++	++++
	Variable	Glu-167-X <sup>§</sup>	128	++++	++++	ND	ND	++++	++++

10 \* The hemagglutination (HA) titer was quantitated after induction with 0.01 mM IPTG. The HA titer represents the greatest bacterial dilution which still agglutinates erythrocytes. HB101/pPAP37 expressing wild type *PapD* has an HA titer of 128.

+ The amount of pili assembled were released from the cells and quantitated after induction with 0.01 mM IPTG. "-" indicates no pili were detected; the degree of piliation of HB101/pPAP37 expressing wild type *PapD* corresponds to "++++".

15 # Chaperone-assisted periplasmic stabilization of pilus subunits *PapA*, *PapE*, *PapF*, *PapK* and *PapG* was determined as described previously (Slonim *et al.*, 1992); "-" indicates no stabilization of the subunit; the number of + 's indicate the degree of subunit stabilization as compared to wild type *PapD*. ND indicates not determined.

§ X indicates mutations to His, Asp, Thr, or Gly.

20 The remarkable consistency of the in vitro and in vivo results and the crystal structure argue that anchoring of pilus subunits by conserved residues Arg-8 and Lys-112 is a critical part of the interaction between *PapD*-like chaperones and pilus biogenesis.

25 Site-directed mutagenesis was performed using the Bio-Rad in vitro mutagenesis kit following the manufacturers instructions. The following primers (non-coding strand) were used to introduce mutations in the *papD* gene (SEQ ID NO: 1):

Thr-7-Val 5'-GTCAAACACCGCCGGAACCTCGTCCAGGCCGA-3' SEQ ID NO: 3

Lys-112-Ala 5'-CGGGCGATAAAAAAGAGCTATTTTGGTCTG-3' SEQ ID NO: 4

Lys-112-Met 5'-GCGATAAAAAAGCATTATTTTCCTCTG-3' SEQ ID NO: 5

Mutations in *papD* were confirmed by sequencing and the altered *papD* genes were cloned into vector pMMB91 under the inducible  $P_{lac}$  promoter as previously described for other *papD* mutations (Slonim et al, 1992). The plasmids were named pThr-7-Val, pLys-112-Ala and pLys-112-Met according to the residue number and mutated amino acid. Plasmid pLS101 is an isogenic construct containing the wild type *papD* gene; plasmid pE167H, pR8G and pR8A contain genes encoding *PapD* with point mutations changing Glu-167 to His and Arg-8 to Gly and Ala, respectively.

The data presented in this example and example 1 demonstrate that the *PapD*-peptide interactions found in the crystal structure reflects the *PapD*-pilus subunit interaction interface in vivo. These interactions define for the first time a mechanism whereby chaperones utilize their immunoglobulin-like domains in a novel recognition paradigm to bind to a diverse group of proteins. *PapD* residues Arg-8 and Lys-112, located in  $\beta$ -strands in the cleft between the two domains, play crucial roles in chaperone recognition by serving as molecular anchors for the carboxyl terminus of pilus proteins. The fact that Arg-8 and Lys-112 are highly conserved amongst periplasmic chaperones indicates that they probably have a universal role in forming the "active site" that binds to the terminal carboxylate group of other protein subunits. Backbone hydrogen bonds along the  $\beta$ -strand of *PapD* subsequently provide strong, sequence-independent interactions along the length of the carboxyl terminus of the subunit. The "zippering" interaction of the conserved alternating hydrophobic residues of the carboxyl terminus of the pilus subunit with the conserved alternating hydrophobic residues in the G1  $\beta$ -strand of the chaperone adds strength and specificity to the binding. Therefore, unlike other immunoglobulin-like proteins (Amit et al, 1988 and de Vos et al, 1992), *PapD* utilizes the  $\beta$ -strand and interdomain cleft

features of its immunoglobulin-like structure to provide a recognition mechanism for binding to several different proteins. Variable residues in loop regions of antibodies provide the exquisite specificity necessary from an extensive  
5 antibody binding repertoire. Similarly, residues in the F1-G1 loop of *PapD* may provide specificity to chaperone binding since these have been found to vary in length and composition amongst members of the periplasmic chaperone family (Holmgren et al, 1992).

### 10 EXAMPLE 3

*Designing of compounds capable of binding to the binding site of PapD*

Having determined the location of a promising binding site for inhibitory ligands as described in Examples 1 and 2, the  
15 computer programs 'PLIM' and 'PLIM\_DBS' (developed by Symbicom AB) were used to find templates for families of compounds capable of binding to the binding site (see the general description above).

PLIM was run on a 20 Å box around the Arg-8 region of the  
20 high-resolution structure of *PapD*, searching for low-energy binding positions for NH, NH<sub>2</sub>, O, OH and C probes. The PLIM runs resulted in a number of suggested positions and orientations of favourable chemical groups (site points) in the region near Arg-8 and Lys-112.

25 Approximately 50 sites were identified, and then triplets and quartets of these site points were selected for their mutual compatibility, that is to say they must have appropriate geometric relationships - distances and orientations that could be found in a real molecule.

30 A search for potential ligands was then made by searching a database for known molecular structures that match the positions of these groups of site points, using PLIM\_DBS.

The best of these database hits were examined visually using a computer-graphic modelling system, and the most promising of these were selected according to a wealth of physico-chemical reasoning.

- 5 Modifications were made as described above to improve binding and simplify synthesis. For example, phenyl rings were added to give better complementarity to the hydrophobic surface of the binding site, charged groups were added to bind to charged groups on the protein that had been outside the scope of Plim, frequently toxic sub-structures were removed, and groups were replaced with functionally similar ones for synthetic reasons. Many such judgements were made, resulting in two families of compounds (hdo and bpy) with differing combinations of characteristics.
- 10
- 15 The efficacy of these modifications was finally assessed using molecular-dynamics free energy calculations as described herein to study the stability of the protein-ligand complex (Åqvist et al., 1994).

- Two structures from one family and one from the second family were then deemed worthy of synthesis and testing.
- 20

The development of a number of the member of these families of compounds will now be described in detail.

- The first family (herein referred to as hdo) was derived from a database entry for 6-hydroxydopamine, denoted hdo\_0 (Fig. 8). The molecule binds to the PapD side chains of Arg-8 and Lys-112, and to the backbone of Lys-110 through hydrogen-bond donation from the hydroxyl groups. The primary-amino group hydrogen-bonds to the side chain of Thr-170.
- 25

- Derivative hdo\_1 (Fig. 8) was created from this base structure by replacing an hydroxyl group with a carboxylic acid that can have a charge interaction with Arg-8 and Lys-112, and replacing the hydroxyl para to this position with a new
- 30

phenyl ring, in order to fill out the hydrophobic cleft in the protein formed by the residues Ile-111, Leu-4 and part of Thr-7. hdo\_4 (Fig. 8) was created from hdo\_1 by replacing the original aromatic ring with a sugar, attaching the hydroxyl, primary amine and carboxylic acid substituents in such a way as to maintain similar relative geometries, and then adding hydroxyl groups to one side of the new benzene ring so as to present a hydrophillic face to the solvent in this region.

At this point the structure looked convincing enough to invest time in a molecular-dynamics simulation, the results of which suggested that the amino group would not in fact bind to the gamma-carbon of Thr-7 of PapD, but would preferentially point into solution, satisfying its hydrogen-bonding requirements with water molecules rather than with the protein side-chain. Thus, the amine group was replaced by the singly donating hydroxyl group, a step that also simplified synthesis, the result being structure hdo\_6 (Fig. 8).

Due to the perceived probability of toxicity due to the hydroxyl substitution pattern on the benzene ring, the para-hydroxyl group was removed to give structure hdo\_7 (Fig. 8), and the remaining hydroxyl group on this ring was replaced by a nitro group for further ease of synthesis (hdo\_8, Fig. 8).

A final compound with the previous nitro group left off was also considered a possibility, and was known as hdo\_9 (Fig. 8).

Another molecular-dynamics simulation was run, on hdo\_9, the results of which suggested that now the predicted interactions would indeed be made and maintained in solution. A statistical analysis of the trajectories with the program "Miss\_Fit" showed that the conformation of the ligand did not change substantially during the run, and changes that did occur were compensated for by complementary shifts in the

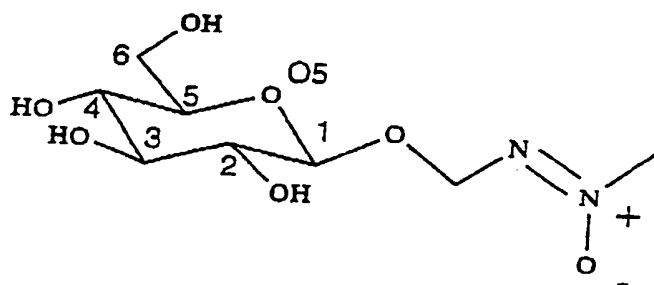


protein structure that maintained all significant interactions.

Dynamics simulation of hdo\_8 was complicated by the lack of availability of suitable molecular-mechanics parameters for the meta-nitro substituted ring, which made assessment through molecular-dynamics simulations harder to rely upon, and this influenced the decision to pursue hdo\_9.

The range of possibilities that were considered during the design phase of the hdo family are summarised in Fig. 9.

- 10 The second family, known as bpy, was derived from a database entry for (methyl-O,N,N-azoxy)-methyl- $\beta$ -D-glucopyranoside cycasin (Kawaminami et al, 1981) denoted as bpy\_0 herein:



bpy\_0

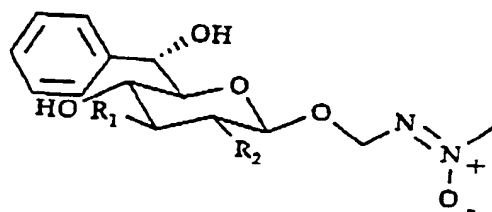
- 15 The conformation stored in the database was non-optimal for binding to the PapD site, though it was close enough to be picked out by the Plim\_DBS search. Small adjustments made to the internal torsion angles improved the binding pattern somewhat.

- 20 It is interesting to note that there are precedents for the approach of using a central carbohydrate as a scaffold to which functionalities are attached as to mimic a peptide (Hirschmann et al, 1992).

Derivative bpy\_1 was then created by adding a phenyl ring to the 6-position, in order to fill out the hydrophobic cleft in

the protein. The 4-OH group should form a favourable dipole interaction with the electron-rich centre of the 6-phenyl group, and together with the 6-OH...O5 intramolecular hydrogen bond should stabilize the desired C5-C6 rotamer.

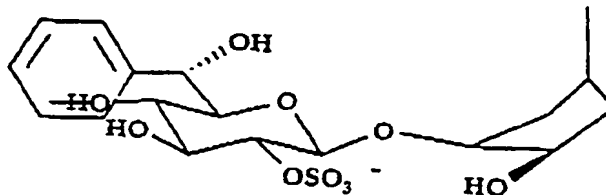
- 5 It was decided that the Lys-112 and Arg-8 side chains that should interact with the hydroxyl group at the 2-position in the ligand do so even more with a charged group, and therefore bpy\_2 was created by replacing this hydroxyl with a sulphate, bpy\_3 by replacing it with a cyclic phosphate, and
- 10 bpy\_5 by replacing it with a non-cyclic phosphate.
- bpy\_4 was created by removing the 3-hydroxyl group from BPy\_2 so that the sulphate group had a better chance of interacting with the Arg-8 side chain, and bpy\_6 was analogously created by removing the 3-hydroxyl group from bpy\_5. The following
- 15 formula shows bpy\_1-bpy\_6:



R <sub>1</sub>	R <sub>2</sub>	
-OH	-OH	bpy_1
-OH	-OSO <sub>2</sub> <sup>-</sup>	bpy_2
-O-PO <sub>2</sub> <sup>-</sup> -O-		bpy_3
-H	-OSO <sub>2</sub> <sup>-</sup>	bpy_4
-OH	-OPO <sub>2</sub> <sup>2-</sup>	bpy_5
-H	-OPO <sub>2</sub> <sup>2-</sup>	bpy_6

Consideration was then given to a small hydrophobic pocket and patches or hydrophobic residues exposed to the solvent near position 1 on the ligand. An aglycon that could make

contact with these parts of the protein and preferably also interact with solvent was sought. These considerations led to the replacement of the azoxy group at position 1 with a 2-hydroxy-4-methylcyclopentyl group, resulting in bpy\_7:

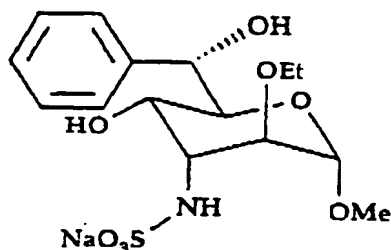


5

bpy\_7

At this point, a molecular dynamics run was made, the results of which suggested that the structure bound well, but that the expected hydrophobic contacts between the cyclopentane and the protein were not maintained.

- 10 Thus, bpy\_9 was created from bpy\_7 with the cyclopentane group removed, replaced by an  $\alpha$ -methyl group for internal stability. Instead, further hydrophobic contact was attempted by the addition of an ethoxy group axial on position 2. The charge was moved from equatorial 2 to axial 3, a move
- 15 stabilised by the inclusion of a NH-link that can donate a hydrogen bond to the oxygen at axial 1. The stereochemical arrangement of functionalities on positions 1-3 now confers conformational rigidity to the structure, as well as a simplification of synthesis.

*bpy\_9*

More dynamics were run, this time suggesting that *bpy\_9* would form a stable complex with the protein, similar to the way that was envisaged during the design.

#### 5 Further developments

By using reasoning in line with the above-described, more compounds have been designed (both from the *hdo*\_family and from the *bpy*\_family). The compounds of interest are those described by the general formulas I and II herein. The most promising compounds resulting from the analyses have been synthesized, cf. example 4.

In the further development of designed compounds using the above-described methods it has interestingly been found that the stability of the binding between the designed compounds and *PapD* seems to be only limitedly dependent on the interaction between the compound and the Arg-8 residue, even though Arg-8 is essential for the *in vivo* function of *PapD*. On the other hand, the stability of the interaction between the designed compounds and *PapD* is still very much dependent on the interaction with Lys-112 as well as with the  $\beta$ -strand. These observations confirm the surprising observation that the stable binding between *PapD* and pilus subunits is very much dependent on the interaction between the  $\beta$ -sheets of the two molecules.

## EXAMPLE 4

*Synthesis of compounds of the hdo family capable of binding to the binding site of PapD*

*General methods*

- 5  $^1\text{H}$ - and  $^{13}\text{C}$ -NMR spectra were recorded in  $\text{CDCl}_3$  at 300 and 75 MHz respectively on a Varian Gemini 300 spectrometer unless otherwise stated. Signals from undeuterated solvent at 7.25 and 77.0 ppm respectively were used as internal reference signals. Optical rotations were measured using a Perkin Elmer  
10 241 polarimeter.

- Thin layer chromatography was performed on Merck DC-Fertig-platten (Kieselgel 60 F<sub>254</sub> 0.25 mm) and spots visualized by spraying with 10% sulphuric acid followed by charring at elevated temperature and/or spraying with molybdato-phosphoric acid hydrate/ $\text{Ce}(\text{SO}_4)_2$ /dil.  $\text{H}_2\text{SO}_4$  followed by heating (molybdatophosphoric acid hydrate =  $\text{H}_3[\text{P}(\text{Mo}_3\text{O}_{10})_4] \times \text{H}_2\text{O}$ , E. Merck, Darmstadt, Germany)  
15

*Ethyl 2,3-O-dibenzoyl-4-O-benzyl-1-thio- $\beta$ -D-glucohexopyranoside*

- 20 Ethyl 2,3-O-dibenzoyl-4,6-O-benzylidene-1-thio- $\beta$ -D-glucohexopyranoside (150 mg, 0.29 mmol) was dissolved together with borane-dimethylamine complex (68 mg, 1.16 mmol) in dry toluene (20 Å). Powdered 4 Å molecular sieves (180 mg) was added and the mixture was stirred at room temperature for 20  
25 minutes. Aluminum trichloride (154 mg, 1.16 mmol) was added and after disappearance of the starting material (approx. 10 min., silica tlc toluene/acetonitrile 4:1) the mixture was filtered, treated with Dowex( $\text{H}^+$ ) ion-exchange resin until the solution was clear, and filtered again. The filtrate was  
30 concentrated and co-evaporated twice with methanol to yield 200 mg residue which upon chromatography on silica (toluene-acetonitrile 8:1) gave 60 mg, 40% of ethyl 2,3-O-dibenzoyl-4-O-benzyl-1-thio- $\beta$ -D-glucohexopyranoside.

<sup>1</sup>H-NMR (CDCl<sub>3</sub>, δ): 8.0-7.2 (m, 15H), 5.75 (t, 9.4 Hz, 1H, H3), 5.37 (t, 9.8 Hz, 1H, H2), 4.74 (d, 9.8 Hz, 1H, H1), 4.61 (s, 2H), 4.02-3.95 and 3.85-3.76 (2 bm, 2H, 2H6) 3.94 (t, 9.5 Hz, 1H, H4), 3.63 (ddd, 9.7 Hz, 3.9 Hz, 2.6 Hz, 1H, H5), 2.74 (2q, 7.5 Hz, 2H), 1.25 (t; 7.5 Hz, 3H).

*Ethyl 6-O-acetyl-2,3-O-dibenzoyl-4-O-benzyl-1-thio-β-D-glucopyranoside*

Ethyl 2,3-O-dibenzoyl-4-O-benzyl-1-thio-β-D-glucopyranoside (200 mg, 0.38 mmol) was dissolved in an ice-cooled mixture of pyridine and acetic anhydride (1:1, 6 Å). The mixture was stirred and attained at room temperature overnight. The mixture was diluted with dichloromethane, washed with satd. aq. sodium hydrogen carbonate and water and dried over sodium sulphate. The solvents were removed under reduced pressure and the residue was subjected to flash chromatography on silica gel (toluene-acetonitrile 6:1) to yield 204 mg, 93% of ethyl 6-O-acetyl-2,3-O-dibenzoyl-4-O-benzyl-1-thio-β-D-glucopyranoside.

<sup>1</sup>H-NMR (CDCl<sub>3</sub>, δ): 8.0-7.2 (m, 15H), 5.75 (t, 9.2 Hz, 1H, H3), 5.40 (t, 9.8 Hz, 1H, H2), 4.71 (d, 10.0 Hz, 1H, H1) 4.59 and 4.55 (2d, 10.8 Hz, 2H, bz1-CH<sub>2</sub>), 4.44 (dd, 12.0 Hz, 1.9 Hz, 1H, H6), 4.27 (dd, 12.2 Hz, 4.6 Hz, 1H, H6') 3.87 (t, 9.5 Hz, 1H, H4), 3.77 (ddd, 9.7 Hz, 4.6 Hz, 1.9 Hz, 1H, H5), 2.73 (m, 2H), 2.09 (s, 3H), 1.26 (t, 3H).

*Methylglycolyl 6-O-acetyl-2,3-O-dibenzoyl-4-O-benzyl-β-D-glucopyranoside*

6-O-Acetyl-2,3-O-dibenzoyl-4-O-benzyl-1-thio-β-D-glucopyranoside (600 mg, 1.04 mmol) and N-iodosuccinimide (468 mg, 2.08 mmol, dried 3h, < 0.1mbar) were dissolved in dry acetonitrile (7 Å, passed through alumina) at room temperature. Methyl glycolate (161 μl, 2.08 mmol) was added and the mixture was stirred for 25 min at room temperature and then cooled on an ice-bath. Trifluoromethylsulphonic acid (18 μl,

0.21 mmol) was carefully added. Complete conversion of starting material had occurred after 10 minutes (tlc: silica toluene-acetonitrile 4:1). The reaction was quenched by addition of triethylamine, the mixture was diluted with  
5 dichloromethane, washed with water, sodium hydrogen carbonate and dried over sodium sulphate. The solvents were removed under reduced pressure and the residue subjected to chromatography on silica (toluene-acetonitrile 11:1) to yield 560 mg, 86% of ethyl 6-O-acetyl-2,3-O-dibenzoyl-4-O-benzyl-1-  
10 thio- $\beta$ -D-glucohexopyranoside.

*Ethyl 2,3-O-Dibenzoyl-4-O-benzyl-1-thio- $\beta$ -D-glucohexopyranoside (hdo\_9:1)*

Ethyl 2,3-O-dibenzoyl-4,6-O-benzylidene-1-thio- $\beta$ -D-glucohexopyranoside (Chemical Abstracts, RN 149563-08-6 or Hällgren and Widmalm, 1993) was dissolved together with borane-dimethylamine complex (68 mg, 1.16 mmol) in dry toluene  
15 (20 ml) (reaction "a" in Fig. 6). Powdered 4 Å molecular sieves (180 mg) were added and the mixture was stirred at room temperature for 20 minutes. Aluminium trichloride (154  
20 mg, 1.16 mmol) was added and after disappearance of the starting material (approx. 10 min., silica tlc toluene/acetonitrile 4:1) the mixture was filtered, treated with Dowex(H<sup>+</sup>) ion-exchange resin until the solution was clear, and filtered again. The filtrate was concentrated and  
25 co-evaporated twice with methanol to yield 200 mg residue which upon chromatography on silica (toluene-acetonitrile 8:1) gave 60 mg, 40% of hdo\_9:1.

<sup>1</sup>H-NMR (CDCl<sub>3</sub>,  $\delta$ ): 8.0-7.2 (m, 15H), 5.75 (t, 9.4Hz, 1H, H3), 5.37 (t, 9.8Hz, 1H, H2), 4.74 (d, 9.8Hz, 1H, H1), 4.61 (s, 2H), 4.02-3.95 and 3.85-3.76 (2 bm, 2H, 2H6) 3.94 (t, 9.5Hz, 1H, H4), 3.63 (ddd, 9.7Hz, 3.9Hz, 2.6Hz, 1H, H5), 2.74 (2q, 7.5Hz, 2H), 1.25 (t; 7.5Hz, 3H).

Ethyl 6-O-acetyl-2,3-O-dibenzoyl-4-O-benzyl-1-thio- $\beta$ -D-glucopyranoside (hdo\_9:2)

Ethyl 2,3-O-dibenzoyl-4-O-benzyl-1-thio- $\beta$ -D-glucopyranoside (hdo\_9:1) (200 mg, 0.38 mmol) was dissolved in an ice-cooled mixture of pyridine and acetic anhydride (1:1, 6 ml). The mixture was stirred and attained room temperature overnight. The mixture was diluted with dichloromethane and washed with saturated aqueous sodium hydrogen carbonate and water, dried over sodium sulphate. the solvents were removed under reduced pressure and the residue was subjected to flash chromatography on silica gel (toluene-acetonitrile 6:1) to yield 204 mg, 93% of hdo\_9:2.

$^1\text{H-NMR}$  ( $\text{CDCl}_3$ ,  $\delta$ ): 8.0-7.2 (m, 15H), 5.75 (t, 9.2Hz, 1H, H3), 5.40 (t, 9.8Hz, 1H, H2), 4.71 (d, 10.0Hz, 1H, H1) 4.59 and 4.55 (2d, 10.8Hz, 2H, bz1-CH2), 4.44 (dd, 12.0Hz, 1.9Hz, 1H, H6), 4.27 (dd, 12.2Hz, 4.6Hz, 1H, H6') 3.87 (t, 9.5Hz, 1H, H4), 3.77 (ddd, 9.7Hz, 4.6Hz, 1.9Hz, 1H, H5), 2.73 (m, 2H), 2.09 (s, 3H), 1.26 (t, 3H).

Methylglycolyl 6-O-acetyl-2,3-O-dibenzoyl-4-O-benzyl- $\beta$ -D-glucopyranoside (hdo\_9:3)

6-O-acetyl-2,3-O-dibenzoyl-4-O-benzyl-1-thio- $\beta$ -D-glucopyranoside (hdo\_9:2) (600 mg, 1.04 mmol) and N-iodosuccinimide (468 mg, 2.08 mmol, dried 3h, < 0.1 mbar) were dissolved in dry acetonitrile (7 ml, passed through alumina) at room temperature. Methyl glycolate (161  $\mu\text{l}$ , 2.08 mmol) was added and the mixture was stirred for 25 min at room temperature and then cooled on an ice-bath. Trifluoromethylsulphonic acid (18  $\mu\text{l}$ , 0.21 mmol) was carefully added. Complete conversion of starting material had occurred after 10 minutes (tlc: silica toluene-acetonitrile 4:1). The reaction was quenched by addition of triethylamine and the mixture was diluted with dichloromethane, washed with water, sodium hydrogen carbonate and dried over sodium sulphate. The solvents were removed under reduced pressure and the residue subjected to chroma-



tography on silica (toluene-acetonitrile 11:1) to yield 560 mg, 86% of (hdo<sub>9</sub>:3).

*Sodium Glycolyl 4-O-benzyl-β-D-glucohexopyranoside (hdo<sub>9</sub>)*

Methylglycolyl 6-O-acetyl-2,3-O-dibenzoyl-4-O-benzyl-β-D-glucohexopyranoside (hdo<sub>9</sub>:3) (350 mg, 0.59 mmol) was dissolved in 7 Å 25 mM sodium methoxide (methanol) and stirred at room temperature for 20 h or until only a very slow-moving spot could be detected on tlc by sulphuric acid charring (silica, toluene-acetonitrile 4:1). A small sample of this solution (=i) was retained as tlc reference before the methylglycolate hydrolysis. Sodium hydroxide (15 mg) and water (200 μl) was added and stirring was continued for 8h or until all "i" was consumed (tlc silica, ethyl acetate-ethanol-acetic acid 8:3:1). The reaction was quenched by bringing the pH to 4.2 with acetic acid. The solvents were removed under reduced pressure and the residue (650 mg) was subjected to flash chromatography on silica (ethyl acetate-ethanol-acetic acid 8:3:1). The material in the R<sub>f</sub>-range 0.15 - 0.25 was collected and concentrated to yield 178 mg, 92% of amorphous white solid.

<sup>1</sup>H-NMR (CD<sub>3</sub>OD, δ relative to methanol at 3.31 ppm): 7.45 (m, 5H), 4.65 (d, 11.0 Hz, 1H), 4.36 (d, 16.0 Hz, 1H), 4.32 (d, 7.9 Hz, 1H, H1), 4.11 (d, 15.8 Hz, 1H), 3.82 (dd, 2.0 Hz, 12.1 Hz, 1H, H6), 3.65 (dd, 4.8 Hz, 12.1 Hz, 1H, H6), 3.59 (t, 9.0 Hz, 1H, H3 or H4), 3.40 (t, 9.1 Hz, 1H, H3 or H4)

This material still contained much acetic acid, so 50 mg was redissolved in water and freeze-dried to give 30 mg of substance free acetic acid but with a water content (<sup>1</sup>H-NMR) of 24.5 mol. eq.

This material was again redissolved in water and 1 equivalent of sodium hydroxide was added, the solution was freeze-dried to give 22 mg of the title compound.

<sup>1</sup>H-NMR (CD<sub>3</sub>OD, δ relative to methanol at 3.31 ppm): 7.40 - 7.22 (m, 5H), 4.95 (d, 10.8 Hz, overlap by residual HDO at 4.93), 4.65 (d, 11.0 Hz, 1H), 4.34 (d, 16.0 Hz, 1H), 4.32 (d, 7.5 Hz, 1H, H1), 4.10 (d, 15.8 Hz), 3.82 (dd, 1.8 Hz, 11.9 Hz, 1H, H6), 3.66 (dd, 4.8 Hz, 12.1 Hz, 1H, H6), 3.59 (t, 6.8 Hz, 1H, H3), 3.41 (t, 9.7 Hz, 1H, H4), 3.15-3.06 (m, 2H, H5, H2).

<sup>13</sup>C-NMR (CD<sub>3</sub>OD, δ relative to methanol at 49.0 ppm): 176.1, 140.0, 129.3, 129.1, 128.6, 104.3, 79.1, 78.1, 77.2, 75.7, 75.2, 68.4, 62.3.

*2-(Hydroxy)ethyl 4-O-benzyl-β-D-glucopyranoside (hdo\_23)*

A mixture of N-Iodosuccinimide (0.109 g, 0.484 mmol) and triflic acid (6.6 μl, 0.0744 mmol) in dichloroethane-diethyl ether (2 ml, 1:1) was added to a stirred mixture of the thioglycoside ethyl 6-O-acetyl-2,3-O-benzoyl-4-O-benzyl-1-thio-β-D-glucopyranoside (hdo\_9:2) (0.210 g, 0.372 mmol), ethylen glycol (0.125 ml, 2.232 mmol), and molecular sieves (0.3 g, 4 Å) in dichloromethane-diethyl ether (3 ml, 1:2) at room temperature during 5 min. After 25 min TLC showed approx. 20 % conversion of the thioglycoside, thus more N-Iodosuccinimide (0.109 g, 0.484 mmol) and triflic acid (6.6 μl, 0.0744 mmol) in dichloroethane-diethyl ether (2 ml, 1:1) was added during 5 min. After additional 30 min. the reaction mixture was filtered through a layer of Celite into an aqueous solution of sodium hydrogen carbonate and sodium bisulphite. The organic layer was separated, washed with water, dried (Na<sub>2</sub>SO<sub>4</sub>) and concentrated. The residue was coevapored twice with toluene, dissolved in methanol containing sodium methoxide (10 ml, 0.2 M), kept 1 h at 50°C, and evaporated. Column chromatography (SiO<sub>2</sub>, chloroform-methanol-water, 100:15:1) of the residue gave hdo\_23 (33 mg, 30 %).

$[\alpha]_D^{22} + 0.9^\circ$  (c 1.4, methanol)

TLC: R<sub>f</sub> 0.29 (chloroform-methanol-water, 100:15:1).

<sup>13</sup>C-NMR (CD<sub>3</sub>OD, δ: relative to methanol at 49.0 ppm) δ:  
139.9, 129.2, 129.0, 128.6, 104.4 (C-1), 79.2, 78.2, 77.0,  
75.7, 75.4, 72.3, 62.3, 62.3 ppm.

<sup>1</sup>H-NMR (CD<sub>3</sub>OD, δ: relative to methanol at 3.31 ppm) δ:  
5 7.45-7.25 (m, 5H), 4.95 (d, 11.0 Hz, 1H, benzylic), 4.65 (d,  
11.0 Hz, 1H, benzylic), 4.29 (d, 7.7 Hz, 1H, H-1) ppm.

#### EXAMPLE 5

*Synthesis of compounds of the bpy family capable of binding  
to the binding site of PapD*

- 10 Methyl 4'-methoxyphenylmethylene-α-D-mannohexopyranosides,  
mixture of 4,6-, 2,3-endo and 2,3-exo monoacetals

(See: Patroni et al, 1988).

- Methyl α-D-mannohexopyranoside (38.8 g, 200 mmol) and p-  
toluenesulphonic acid (200 mg) were dissolved in dimethylfor-  
15 mamide (300 ml) and stirred at 100°C under a nitrogen stream.  
4-methoxy-benzaldehyde dimethyl acetal (37.7 g, 240 mmol) in  
dimethylformamide (300 ml) was added dropwise under 3h. Since  
starting material still remained (tlc, silica, ethyl ace-  
20 tate), stirring and heating was continued for 2h. The compo-  
sition of the reaction mixture did not show any visible  
change at that time (tlc, silica, ethyl acetate) and the  
reaction was quenched with potassium carbonate (2.15 g). The  
solvent was removed under reduced pressure and the residue  
was filtered through silica (ethyl acetate) to remove salts  
25 and starting material. The concentrated eluents deposited  
monoacetals (<sup>1</sup>H-NMR) on trituration with methyl t-butyl ether  
(11 g, 18%), which was collected by filtration. The filtrates  
were subjected to flash chromatography to yield additional 20  
g of the monoacetal mixture (yield 51%) and 17 g of the  
30 2,3;4,6-di acetals (endo/exo = 1:1).

Methyl 2-O-ethyl-4,6-O-(4'-methoxy)phenylmethylene- $\alpha$ -D-mannohexopyranoside (bpy\_9:1)

Methyl 4'-methoxyphenylmethylene- $\alpha$ -D-mannohexopyranosides, mixture of 4,6-, 2,3-endo and 2,3-exo monophenylmethylene acetals (11.3 g, 36.2 mmol), ethyl iodide (4.35 ml, 54.3 mmol) and tetrabutylammonium hydrogen sulphate (2.48 g, 7.24 mmol) were dissolved in methylene chloride (500 ml). To this was added a solution of sodium hydroxide (2.5 g, 140 mmol) in water (50 ml) and the mixture was heated at reflux for four days, with daily additions of ethyl iodide (2.4 ml) and 20% sodium hydroxide (1 ml). After this time, no visible change in the composition of the reaction mixture could be detected by tlc (toluene-ethyl acetate 2:1). The mixture was allowed to cool, the phases separated and the organic phase washed with water (3x) and dried over sodium sulphate. The solvent was removed under reduced pressure and the residue was subjected to flash chromatography on silica gel (toluene-ethyl acetate 6:1). The fractions containing the component with  $R_f = 0.40$  on tlc (toluene-ethyl acetate 2:1) were pooled and rechromatographed as above to yield 3.15 g, 26% of the title compound. Fractions containing components with  $R_f = 0.32$  on tlc were found to consist of a mixture of monoethylated monophenylmethylene acetal regio-isomers ( $^1\text{H-NMR}$ ).

$$[\alpha]_D^{20} = +11.2^\circ (\text{c}1.02, \text{CHCl}_3)$$

$^1\text{H-NMR}$  ( $\text{CDCl}_3$ ,  $\delta$ ): 7.42 (symm.m, 2H), 6.88 (symm.m, 2H), 5.53 (s, 1H), 4.77 (d, 1.3 Hz, 1H, H1), 4.24 (dd, 4.2 Hz, 9.7 Hz, 1H, H6), 4.02 (dt, 4.0 Hz, 8.8 Hz, 1H, H5), 3.83 (dd, 1.8 Hz, 9.7 Hz, 1H, H6), 3.80-3.69 (s and overl.m, 6H, 1-OMe, H3, H4, 2-O-CH $\text{HCH}_3$ ), 3.66 (dd, 1.3 Hz, 3.7 Hz, 1H, H2), 3.64-3.55 (m, 1H, 2-O-CH $\text{HCH}_3$ ), 3.69 (s, 3H, Ar-OMe), 2.48 (d, 1H, OH), 1.26 (t, 7.0 Hz, 3H, 2-O-CH $_2\text{CH}_3$ ).

$^1\text{H-NMR}$  ( $\text{CDCl}_3$ , with added  $\text{Cl}_3\text{CCONC}$ ,  $\delta$ ): 8.54 (s, 1H, acylcarbamate NH), 7.38 (symm.m, 2H), 6.88 (symm.m, 2H), 5.52 (s, 1H), 5.27 (dd, 3.5 Hz, 10.3 Hz, 1H, H3), 4.76 (d, 1.6 Hz, 1H,

H1), 4.32-4.20 (m, 1H), 4.13 (symm.m, 1H, H5), 3.97 (dd, 1.8 Hz, 3.5 Hz, 1H, H2), 3.91-3.80 (m, 1H), 3.79 (s, 3H, 1-OMe), 3.77-3.54 (m, 3H), 3.41 (s, 3H, Ar-OMe), 1.23 (t, 7.0 Hz, 3H).

5  $^{13}\text{C}$ -NMR ( $\text{CDCl}_3$ ,  $\delta$ ): 160.1, 129.9, 127.6, 113.6, 102.0, 99.3, 79.4, 78.8, 68.7, 68.3, 67.2, 63.2, 55.2, 54.9, 15.4.

Calc. for  $\text{C}_{17}\text{H}_{25}\text{O}_7$ : C: 59.8 H: 7.38,

*Methyl 2-O-ethyl-3-O-dimethyl-t-butylsilyl-4,6-O-(4'-methoxy)phenylmethylen- $\alpha$ -D-mannohexopyranoside (bpy\_9:2)*

10 Methyl 2-O-ethyl-4,6-O-(4'-methoxy)phenylmethylen- $\alpha$ -D-mannohexopyranoside (bpy\_9:1) (0.98 g, 2.88 mmol) was dissolved in dry methylene chloride (18 ml) together with triethyl amine (600  $\mu\text{l}$ , 4.32 mmol) and the mixture was ice-cooled. t-Butyldimethylsilyl trifluoromethyl sulphonate (795  $\mu\text{l}$ , 3.45 mmol)  
15 in methylene chloride (2 ml) was added dropwise and the reaction mixture with the cooling bath was left at ambient temperature overnight (tlc silica toluene-methyl t-butyl ether 4:1). The excess silylating reagent was destroyed by addition of methanol (3 ml), the mixture was concentrated  
20 under reduced pressure and the residue was subjected to flash chromatography on silica (toluene-methyl t-butyl ether 40:1) to yield 1.40 g, 100% of (bpy\_9:2).

$^1\text{H}$ -NMR ( $\text{CDCl}_3$ ,  $\delta$ ): 7.40 (symm.m, 2H), 6.87 (symm.m, 2H), 5.53 (s, 1H), 4.70 (d, 1.3 Hz, 1H, H1), 4.21 (dd, 4.6 Hz, 9.9 Hz, 1H, H6), 4.10 (dd, 3.3 Hz, 9.7 Hz, 1H, H3), 3.91 (t, 9.5 Hz, 1H, H4), 3.90-3.79 (s overl. m, 5H, H6, 2-OCH<sub>2</sub>HCH<sub>3</sub>; 1-OMe), 3.76-3.63 (m, 2H, H5, 2-OCH<sub>2</sub>HCH<sub>3</sub>), 3.55 (dd, 1.5 Hz, 3.3 Hz, 1H, H2), 3.67 (s, 3H, Ar-OMe), 1.24 (t, 7.0 Hz, 3H), 0.88 (s, 9H), 0.08 (s, 3H), 0.03 (s, 3H).

30  $^{13}\text{C}$ -NMR ( $\text{CDCl}_3$ ,  $\delta$ ): 159.9, 130.3, 127.5, 113.4, 101.8, 101.2, 79.9, 79.1, 70.3, 68.8, 68.2, 64.1, 55.2, 54.8, 25.8, 18.3, 15.6, -4.4, -4.9.

Methyl 2-O-ethyl-3-O-dimethyl-t-butyldisilyl-4, -O-(4'-methoxy)-benzyl- $\alpha$ -D-mannohexopyranoside (bpy\_9:3)

and

5 methyl 2-O-ethyl-3-O-dimethyl-t-butyldisilyl-6-O-(4'-methoxy)-benzyl- $\alpha$ -D-mannohexopyranoside (bpy\_9:8)

Methyl 2-O-ethyl-3-O-dimethyl-t-butyldisilyl-4,6-O-(4'-methoxy)phenylmethylen- $\alpha$ -D-mannohexopyranoside (bpy\_9:2) (3.61 g, 7.95 mmol) was dissolved in acetonitrile (150 ml, dried over 4 Å molecular sieves), powdered 3 Å molecular sieves (6 g) was added and the mixture cooled to -30°C under a nitrogen atmosphere and stirred for 15 min. Sodium cyanoborohydride (3.04 g, 47.7 mmol) was added and a solution of chlorotrimethylsilane (7.95 ml, 47.7 mmol) in dry acetonitrile (50 ml, dried over 4 Å molecular sieves) was added dropwise in 1h. 15 Stirring was continued for 2h at -30°C, whereafter the temperature was allowed to rise to 0°C. After 1h tlc (toluene-ethyl acetate 4:1) showed complete conversion of starting material. The mixture was filtered through a Celite® pad, the Celite® washed thoroughly with ethyl acetate. The combined 20 filtrates were washed with aq. satd. sodium hydrogen carbonate, aq. satd. sodium chloride, concentrated under reduced pressure, redissolved in toluene and concentrated under reduced pressure. The residue was subjected to flash chromatography on silica (toluene-ethyl acetate 4:1, then 1:1) to 25 obtain bpy\_9:8 (= bpy\_21:1) 578 mg, 16% and (bpy\_9:3) 2.34 g, 65%.

bpy\_9:3 (some assignments from COSY, HETCOR):

<sup>1</sup>H-NMR (CDCl<sub>3</sub>,  $\delta$ ): 7.25 (symm.m 2H), 6.87 (symm.m, 2H), 4.61 (d, 11.0 Hz, 1H), 4.66 (d, 1.8 Hz, 1H, H1), 4.51 (d, 10.8 Hz, 1H), 4.02 (dd, 3.1 Hz, 5.9 Hz, 1H, H3), 3.80 (s, 3H, 1-OMe), 3.79-3.60 (m, 5H, H4, 2H6, 2-OCH<sub>2</sub>CH<sub>3</sub>), 3.54 (ddd, 2.9 Hz, 4.8 Hz, 9.7 Hz, 1H, H5), 3.46 (dd, 2.0 Hz, 3.3 Hz, 1H, H2), 3.33 (s, 3H, Ar-OMe), 2.10 and 1.86 (2 v.br.s, -1H, 6-OH), 1.23 (t, 6.8 Hz, 3H), 0.95 (s, 9H), 0.12 (s 6H).

113

<sup>1</sup>H-NMR (CDCl<sub>3</sub>, with added Cl<sub>3</sub>CCONC, δ): 8.40 (s, 1H, acylcarbamate NH), 7.24 (symm.m, 2H, Ar H3',5'), 6.86 (symm.m, 2H, Ar H2',H6'), 4.82 (d, 11.2 Hz, 1H), 4.67 (d, 2.0 Hz, 1H, H1), 4.43-4.32 (m, 2H, 2H6), 4.08-3.98 (m, 1H, H3), 3.78 (s, 3H, 1-OMe), 3.77-3.61 (m, 4H, H5, 2-OCH<sub>2</sub>CH<sub>3</sub>, H4), 3.46 (dd, 1.8 Hz, 3.1 Hz, 1H, H2), 3.34 (s, 3H, Ar-OMe), 1.22 (t, 7.0 Hz, 3H), 0.96 (s, 9H), 0.14 (s, 3H), 0.13 (s, 3H).

<sup>13</sup>C-NMR (CDCl<sub>3</sub>, δ): 159.1, 130.7, 129.3, 113.8, 99.8, 79.5, 75.7, 74.7, 73.0, 72.2, 67.4, 62.5, 55.2, 54.7, 18.0, 15.7, 4.3, -4.6.

<sup>13</sup>C-NMR (CDCl<sub>3</sub>, with added Cl<sub>3</sub>CCONC, δ): 159.3 (Ar C4'), 157.4, 149.3, 130.2 (Ar C2'), 129.6 (Ar C3',C5'), 113.8 (Ar C2',C6'), 99.6 (C1), 79.2 (C2), 74.56\* (4-OCH<sub>2</sub>Ar or 2-CH<sub>2</sub>CH<sub>3</sub>), 74.54\* (4-OCH<sub>2</sub>Ar or 2-CH<sub>2</sub>CH<sub>3</sub>), 73.2 (C3), 69.7 (C5), 67.2 (C4), 66.3 (C6), 55.2 (1-OMe), 54.9 (Ar-OMe), 25.9 (CMe<sub>3</sub>), 18.0 (CMe<sub>3</sub>), 15.6, -4.3, -4.7.

bpy<sub>9</sub>:8:

<sup>1</sup>H-NMR (CDCl<sub>3</sub>, δ): 7.28 (symm.m, 2H), 6.86 (symm.m, 2H), 4.70 (d, 1.8 Hz, 1H, H1), 4.55 (d, 11.7 Hz, 1H, H6), 4.51 (d, 11.6 Hz, 1H, H6), 3.89-3.55 and 3.79 (m and s, 7H, H3, H4, H5, 2-OCH<sub>2</sub>CH<sub>3</sub>, 1-OMe), 3.44 (dd, 1.8 Hz, 2.9 Hz, 1H, H2), 3.56 (s, 3H, Ar-OMe), 2.29 (d, 2.0 Hz, 1H, 4-OH), 1.20 (t, 7.0 Hz, 3H), 0.92 (s, 9H), 0.13 (s, 3H), 0.12 (s, 3H).

<sup>13</sup>C-NMR (CDCl<sub>3</sub>, δ): 159.1, 130.3, 129.3, 113.7, 99.7, 78.9, 73.19\*, 73.16\*, 71.3, 70.3, 69.1, 67.2, 55.2, 54.8, 25.8, 18.2, 15.6, -4.5, -4.6.

\*resolved by multiplying the FID with a gaussian weight-function.

Methyl 2-O-ethyl-3-O-dimethyl-t-butyldisilyl-4,-O-(4'-methoxy)-benzyl-6(S)-phenyl- $\alpha$ -D-mannohexopyranoside (bpy\_9:4)

(Oxidation ref. D.F. Taber et al, J. Org. Chem. 52, 5621-2, 1987) Methyl 2-O-ethyl-3-O-dimethyl-t-butyldisilyl-4,-O-(4'-methoxy)benzyl- $\alpha$ -D-mannohexopyranoside (bpy\_9:3) (200 mg, 0.438 mmol) and dimethylsulphoxide (64.0  $\mu$ l, 0.876 mmol) were dissolved in dichloromethane (2.5 ml, dried over 4 Å molecular sieves) and the solution was cooled on an ice-bath. Phosphorpentoxide (124 mg, 0.876 mmol) was added quickly and the resulting stirred suspension was allowed to attain RT (1h). The mixture was again ice-cooled and triethyl amine (215  $\mu$ l, 1.53 mmol) was added causing an immediate dissolution of the gel-like suspension. The ice-bath was removed and the mixture was stirred at RT for 1h, the diluted with ethyl acetate, washed with aq. 0.2 M sodium dihydrogen phosphate, dried over sodium sulphate and concentrated under reduced pressure (< 0.02 mbar). The residue was dissolved in tetrahydrofuran (2 ml, dried over 4 Å molecular sieves), stirred and cooled to -20°C. Phenylmagnesium chloride in tetrahydrofuran (345  $\mu$ l of a 25 w% solution, 0.657 mmol, Janssen, Belgium) was added dropwise (3 min). The reaction was monitored with tlc (toluene-ethyl acetate 4:1) and quenched with aq. 10% ammonium sulphate (3 ml). The mixture diluted with ethyl acetate and washed twice with water, dried over sodium sulphate and concentrated under reduced pressure. The residue was subjected to flash chromatography on silica (toluene-methyl t-butyl ether 20:1). The fractions containing the component with a  $R_f$  = 0.44 on tlc (toluene-ethyl acetate 4:1) were pooled and concentrated under reduced pressure to yield 65 mg, 30% of the title compound.

Molecular mechanics calculations with MM2(91) (Burkert and Allinger, 1982) of the R and S diastereomers and subsequent calculations with Karplus equation (Bothner-By, 1965) predicted that the  $^1\text{H-NMR}$   $J_{5,6}$  would be 4.8 and 6.2 respectively. This difference is too small to be but an indication of the stereochemistry. Furthermore, the H6 signal is poorly



resolved and the H5 signal is severely overlapped by other resonances. The  $^1\text{H}$ -COSY H4-H5 crosspeak, however, contains the sum  $J_{4.5} + J_{5.6}$  and since the  $J_{4.5}$  is observable in the H4-resonance, the  $J_{5.6}$  can be assigned the value 6.8 Hz. This indicates that the isolated product has the 6-(S) stereochemistry.

After the removal of the 4-methoxybenzyl group the calculated coupling constants become 5.5 Hz and 12.9 Hz, respectively.

$^1\text{H}$ -NMR ( $\text{CDCl}_3$ ,  $\delta$ ): 7.44-7.19 (m, 7H), 6.89 (symm.m, 2H), 5.00 (br.m, 1H, H6), 4.90 (d, 10.8 Hz, 1H), 4.66 (d, 10.8 Hz, 1H), 4.58 (d, 1.8 Hz, 1H, H1), 4.04 (dd, 3.1 Hz, 9.0 Hz), 3.99 (t, 9.2 Hz, 1H, H4), 3.81 and 3.82-3.56 (s and m, 6H, 1-OMe, 2- $\text{CH}_2\text{CH}_3$ ), 3.44 (dd, 2.0 Hz, 2.9 Hz, 1H, H2), 2.84 (s, 3H, Ar-OMe), 1.24 (t, 3H), 1.97 (s, 9H), 0.15 (s, 3H), 0.13 (s, 3H).

$^{13}\text{C}$ -NMR ( $\text{CDCl}_3$ ,  $\delta$ ): 159.0, 142.5, 130.8, 129.2, 127.7, 126.7, 126.6, 125.7, 113.7, 99.5, 79.8, 76.1, 74.5, 73.0, 70.9, 67.9, 55.1, 54.1, 25.8, 17.9, 15.4, -4.5, -4.8.

Methyl 2,3-anhydro-4,6-O-p-methoxybenzylidene- $\alpha$ -D-mannopyranoside (bpy\_9:9)

To a stirred suspension of sodium hydride (2.6 g, 65 mmol, 65% dispersion in mineral oil) in N,N-dimethylformamide (150 ml) was added with stirring a solution of methyl 4,6-O-p-methoxybenzylidene- $\alpha$ -D-glucopyranoside (9.36 g, 30 mmol) in N,N-dimethylformamide (65 ml). The mixture was stirred to 45 minutes and p-toluenesulfonylimidazol (7.24 g, 33 mmol) was added. Stirring was continued for 2 hours and the mixture was the poured into ice-water. The precipitate was filtered of, dried in vacuo to give crude bpy\_9:) (7.7 g).

Recrystallization twice from methanol (250 ml gave bpy\_9:9 (3.58 g). Chromatography ( $\text{SiO}_2$ , ethyl acetate-heptane, 2:3) of the mother liquor followed by crystallization from me-

116

thanol (150 ml) gave additional bpy\_9:9 (1.87 g). The total yield of bpy\_9:9 was 5.45 g (61%).

m.p. 152-153.5°C (methanol)

$[\alpha]_D^{22}$ : 96.4° (c 1.8, chloroform)

- 5  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$ : 7.43 and 6.91 (AB pattern, further coupled, 2H,  $J_{AB}=8.7$  Hz), 5.53 (s, 1H, ArCH), 4.90 (s, 1H, H-1), 4.30-4.19 (sym.m, 1H) 3.82 (s, 3H,  $\text{CH}_3\text{OAr}$ ), 3.78-3.64 (3H), 3.49-3.46 (4H), 3.48 (s, 3H,  $\text{CH}_3\text{O}$ ), 3.18 (d, 1H,  $J$  3.7=Hz, H-2/H-3) ppm.
- 10  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ )  $\delta$ : 160.3, 129.6, 127.5, 113.7, 102.4, 96.9, 74.8, 69.4, 61.7, 55.7, 55.3, 53.8 and 50.5 ppm.

Methyl 3-azido-4,6-O-p-methoxybenzylidene- $\alpha$ -D-altropyranoside (bpy\_9:10)

- 15 Methyl 2,3-anhydro-4,6-O-p-methoxybenzylidene- $\alpha$ -D-mannopyranoside (bpy\_9:9, 2.35 g, 8 mmol), sodium azide (2.08 g, 32 mmol) and ammonium chloride (0.86 g, 16 mmol) was stirred for 5 hours at 110°C in a mixture of 2-methoxyethanol (25 ml) and water (5 ml). The suspension gradually dissolved to give a turbid solution. The mixture was then partitioned between
- 20 ethyl acetate and aqueous NaOH (0.25 M, 100 ml). The aqueous phase was extracted twice with ethyl acetate and the combined organic phases was washed twice with water and then with saturated aqueous NaCl, dried ( $\text{Na}_2\text{SO}_4$ ), filtered and concentrated. Chromatography ( $\text{SiO}_2$ ) ethyl acetate-heptane, 2:3-2:1)
- 25 gave bpy\_9:10 (1.97 g, 73%). An analytical sample was crystallized from ethyl acetate-heptane.

m.p. 91-93°C (EtOAc/Heptane)

$[\alpha]_D^{22}$ : +26.1° (c 1.1, chloroform)

- 30  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$ : 7.44 and 6.91 (AB pattern further coupled, 2H,  $J_{AB}=8.6$  Hz), 5.57 (s, 1H, ArCH), 4.56 (s, 1H, H-1), 4.43-4.21 (2H), 4.16-4.08 (1H), 4.03 (unresolved dd, 1H,  $J_1=J_2=3$

117

Hz), 3.91 (bs, 1H, H-2), 3.83-3.74 (4H), 3.80 (s, 3H, CH<sub>3</sub>O), 3.43 (s, 3H, CH<sub>3</sub>O), 2.36 (bs, 1H, OH) ppm.

<sup>13</sup>C NMR (CDCl<sub>3</sub>) δ: 160.3, 129.5, 127.5, 113.7, 102.3, 101.4, 75.8, 69.8, 69.0, 60.1, 59.0, 55.7 and 55.3 ppm.

- 5 Acetylation (acetic anhydride-pyridine, 3:5) of a sample of bpy\_9:10 gave an acetate which had the following <sup>1</sup>H NMR spectrum (assignments were confirmed by the corresponding homonuclear COSY spectrum):

10 <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ: 7.43 and 6.91 (AB pattern further coupled, 2H, J<sub>AB</sub>=8.8 Hz), 5.59 (s, 1H, ArCH), 4.97 (dd, 1H, J=2.2 and 0.9 Hz, H-2), 4.57 (s, 1H, H-1), 4.36 (2H, H-6 and H-4), 4.09-4.01 (2H, H-6, and H-3), 3.84-3.74 (2H, CH<sub>3</sub>O and H-5), 3.81 (s, 3H, CH<sub>3</sub>O), 3.44 (s, 3H, CH<sub>3</sub>O), 2.15 (s, 3H, CH<sub>3</sub>CO) ppm.

- 15 Methyl 3-azido-2-O-ethyl-4,6-O-p-methoxybenzylidene-α-D-altropyranoside (bpy\_9:11)

Methyl 3-azido-4,6-O-p-methoxybenzylidene-α-D-altropyranoside (bpy\_9:10, 1.7 g, 5 mmol), barium oxide (3.0 g, 19.6 mmol) and ethyl iodide (5 ml, 62 mmol) was stirred in dimethyl sulphoxide (5 ml). Water (10 μl) was added and the stirring was continued for 16 hours. The mixture was then partitioned between ethyl acetate and water. The aqueous phase was extracted twice with ethyl acetate and the combined organic phases was washed twice with water and then with saturated aqueous NaCl, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and concentrated. 20  
25 Chromatography (SiO<sub>2</sub>, ethyl acetate-heptane, 1:3) gave bpy\_9:11 as an oil (1.7 g, 92%).

[α]<sub>D</sub><sup>22</sup>: 19.1° (c 1.6, chloroform).

30 <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ: 7.44 and 6.91 (AB pattern, further coupled, 2H, J<sub>AB</sub>=8.6 Hz), 5.57 (s, 2H, ArCH), 4.63 (bs, 1H, H-1).

(sym.m.2H), 3.52 (dd, 1H,  $J=2.4$  and  $0.9$  Hz, H-3), 3.34 (s, 3H,  $\text{CH}_3\text{O}$ ), 1.25 (t, 3H,  $J=6.9$  Hz,  $\text{CH}_3\text{CH}_2$ ) ppm.

$^{13}\text{C}$  NMR ( $\text{CDCl}_3$ )  $\delta$ : 160.1, 129.6, 127.4, 113.6, 102.1, 99.6, 77.0, 76.1, 69.0, 66.5, 58.7, 58.1, 55.5, 55.2 and 15.3 ppm.

5 Methyl 3-azido-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl- $\alpha$ -D-altro-pyranoside

(Cf. R. Johansson and B. Samuelsson; *J. Chem. Soc. Commun.* 1984, 201-202).

Methyl 3-azido-3-deoxy-2-O-ethyl-4,6-O-p-methoxybenzylidene-  
10  $\alpha$ -D-altropyranoside (bpy<sub>9</sub>:11) (2.41 g, 6.6 mmol), sodium cyanoborohydride (2.5 g, 39.6 mmol) and ground molecular sieves (5 g) was stirred in acetonitrile (130 ml) at  $0^\circ\text{C}$ . A  
15 solution of trimethylsilyl chloride (5.0 ml, 40 mmol) in acetonitrile (45 ml) was added during 55 min. After additional 1.5 h of stirring, the cooling bath was removed and the stirring was continued for 21 h. The mixture was filtered through celite and the filtrate was partitioned between ethyl acetate and saturated aqueous  $\text{NaHCO}_3$ . The organic phase was  
20 washed twice with saturated aqueous  $\text{NaHCO}_3$  and then with saturated aqueous  $\text{NaCl}$ , dried ( $\text{Na}_2\text{SO}_4+\text{NaHCO}_3$ ), filtered and evaporated. Chromatography twice ( $\text{SiO}_2$ , ethyl acetate-heptane 2:7+2:1 and ethyl acetate-toluene 1:2+2:1) gave the title compound (2.23 g, 92 %) as an oil.

$[\alpha]_D^{22}$ :  $+113.6^\circ$  ( $c$  1.4, chloroform)

25  $^1\text{H}$  NMR data ( $\text{CDCl}_3$ )  $\delta$ : 7.30 and 6.89 (AB pattern, further coupled, 4 H,  $J = 8.8$  Hz), 4.62 and 4.56 (AB pattern, 2 H,  $J = 11.3$  Hz, benzylic H), 4.58 (unresolved d, virtually coupled, 1 H,  $J < 2$  Hz, H-1), 4.00-3.84 (3 H, H-3, H-4 and H-5), 3.81 (s, 3 H,  $\text{CH}_3\text{OAr}$ ), 3.84-3.7 (2 H, H-6), 3.62-3.43 (3  
30 H,  $\text{CH}_2\text{CH}_3$  and H-2), 3.38 (s, 3 H,  $\text{CH}_3\text{O}$ ), 1.99 (bs, 1 H, OH) and 1.19 (t, 3 H,  $J = 7.0$  Hz,  $\text{CH}_2\text{CH}_3$ ) ppm.

<sup>13</sup>C NMR data (CDCl<sub>3</sub>) δ: 159.6, 129.9, 129.6, 113.9, 99.9, 76.9, 72.2, 71.9, 67.7, 66.3, 62.4, 58.6, 55.4, 55.3 and 15.4 ppm.

Methyl 3-azido-6-O-benzoyl-3-deoxy-2-O-ethyl-4-O-p-methoxy-  
5 benzyl-α-D-altropyranoside

A solution of methyl 3-azido-3-deoxy-2-O-ethyl-4-O-p-methoxy-  
benzyl-α-D-altro-pyranoside (1.47 g, 4 mmol) in pyridine (35  
ml) was stirred at 0°C. Benzoyl chloride (1.2 ml, 10.3 mmol)  
was added and the temperature was raised to room temperature.  
10 The stirring was continued for 1.5 h and the mixture was then  
cooled to 0°C. Methanol (10 ml) was added and the mixture was  
stirred for additional 20 min. at room temperature, evapo-  
rated and partitioned between ethyl acetate and water. The  
organic phase was washed subsequently with water, saturated  
15 aqueous NaHCO<sub>3</sub>, water, saturated aqueous NaCl, dried  
(Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated. Chromatography (SiO<sub>2</sub>,  
ethyl acetate-heptane 1:4) gave the benzoate (1.64 g, 87%).

[α]<sub>D</sub><sup>22</sup>: +0.95° (c 1.0, chloroform)

<sup>1</sup>H NMR data (CDCl<sub>3</sub>) δ: 8.04-7.98 (m, 2 H), 7.56 (tt, 1 H, J =  
20 7.5 and 1.3 Hz), 7.43 (t, further coupled, 2 H, J = 7.5 Hz),  
7.26 and 6.83 (AB pattern, further coupled, 4 H, J = 8.6 Hz),  
4.62 and 4.50 (AB pattern, 2 H, J = 11.3 Hz, ArCH<sub>2</sub>), 4.63 (d,  
further coupled, J = 1.3 Hz, H-1), 4.54 (A part of an ABX  
system, 1 H, J = 11.8 and 2.7 Hz, H-6<sub>a</sub>), 4.47 (B part of an  
25 ABX system, 1 H, J = 11.8 and 5.5 Hz, H-6<sub>b</sub>), 4.27 (ddd, 1 H,  
J = 8.4, 5.5 and 2.7 Hz, H-5), 3.99-3.92 (2 H, H-3 and H-4),  
3.74 (s, 3 H, CH<sub>3</sub>OAr), 3.67-3.48 (3 H, CH<sub>2</sub>CH<sub>3</sub> and H-2) and  
1.22 (t, 3 H, J = 7.0 Hz, CH<sub>2</sub>CH<sub>3</sub>) ppm.

<sup>13</sup>C NMR data (CDCl<sub>3</sub>) δ: 166.2, 159.5, 132.9, 130.1, 129.9,  
30 129.6, 129.2, 128.3, 113.9, 99.6, 76.72, 72.0, 71.7, 62.3,  
62.2, 64.1, 58.5, 55.4, 55.2 and 15.4 ppm.

*Methyl 6-O-benzoyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl-3-sulfamino- $\alpha$ -D-altropyranoside sodium salt*

(Cf. H. P. Wessel; *J. Carbohydr. Chem.* 11 (8), (1992), 1039-1052).

- 5 To a stirred solution of methyl 3-azido-6-O-benzoyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl- $\alpha$ -D-altropyranoside (620 mg, 1.31 mmol) in tetrahydrofuran (40 ml) was added water (265 ml) and triphenylphosphine (1.7 g, 6.5 mmol). The solution was stirred for 24 h and was then concentrated to 12 ml. The  
10 residue was diluted with methanol (10 ml) and water (3 ml). To the stirred solution was then added trimethylamine sulfurtrioxide complex (370 mg, 2.7 mmol). The pH was adjusted to 8.5 by 1 M aqueous NaOH. After a few minutes a precipitate was formed, which was dissolved by adding tetrahydrofuran (3  
15 ml) and methanol (4 ml). The reaction mixture was stirred for 0.5 h, after which time additional trimethylamine sulfurtrioxide complex (86 mg) was added. The stirring was continued for 15 min. During the course of the reaction pH was kept at 8-9 by adding 1 M aqueous NaOH. The mixture was then diluted  
20 with water (7 ml) and the organic solvents were evaporated under reduced pressure. Additional water (5 ml) was added and the suspension was lyophilised and chromatographed (SiO<sub>2</sub>, chloroform-methanol-water 100:15:1→70:30:5) to give the sulfamino compound, presumably in the sodium salt form, (668  
25 mg, 92 %) after lyophilisation from water. A sample was crystallized from aqueous acetone.

m.p. 112-115°C (aqueous acetone).

$[\alpha]_D^{22}$ : +156.4° (c 0.5, water)

- <sup>1</sup>H NMR data (CD<sub>3</sub>OD)  $\delta$ : 7.94-7.89 (m, 2 H), 2 H, 7.64-7.56 (m, 1 H), 7.44 (m, 2 H), 7.28 and 6.73 (AB pattern, further  
30 coupled, 4 H, J = 8.6 Hz), 4.71 and 4.49 (AB pattern, 2 H, J = 11.4 Hz, ArCH<sub>2</sub>), 4.68 (bs, 1 H, H-1), 4.51 (dd, 1 H, J = 11.6 and 2.4 Hz, H-6<sub>a</sub>), 4.44 (dd, 1H, J = 11.6 and 4.4 Hz, H-6<sub>b</sub>), 4.09 (unresolved ddd, 1 H, J = 4, 3 and 1 Hz, H-3),

121

4.00-3.92 (2 H, H-5 and H-2), 3.86 (dd, 1 H,  $J = 9.9$  and  $4.2$  Hz, H-4), 3.71-3.54 (5 H,  $\text{CH}_2\text{CH}_3$  and  $\text{CH}_3\text{OAr}$ ), 3.65 (s,  $\text{CH}_3\text{OAr}$ ), 3.40 (s, 3 H,  $\text{CH}_3\text{O}$ ) and 1.17 (t, 3 H,  $J = 7.0$  Hz,  $\text{CH}_2\text{CH}_3$ ) ppm.

- 5  $^{13}\text{C}$  NMR data ( $\text{CD}_3\text{OD}$ )  $\delta$ : 167.8, 160.8, 134.2, 131.4, 130.9, 130.6, 129.5, 114.7, 101.6, 77.03, 70.02, 69.2, 67.0, 66.4, 65.2, 55.6, 55.5, 50.6 and 15.8 ppm.

*Methyl 6-O-benzoyl-3-deoxy-2-O-ethyl-3-sulfamino- $\alpha$ -D-altropyranoside ammonium salt (bpy\_30)*

- 10 Methyl 6-O-benzoyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl-3-sulfamino- $\alpha$ -D-altro-pyranoside sodium salt (334 mg, 2.0 mmol) was hydrogenolysed for 6 h in glacial acetic acid (55 ml) at 0.23 MPa and room temperature, using 5 % palladium on charcoal as a catalyst. Filtering and lyophilisation gave a  
15 residue which was chromatographed ( $\text{SiO}_2$ , chloroform-methanol-water 100:15:1 $\rightarrow$ 70:30:5). The product was dissolved in water and passed through a cation exchange resin (BIO-REX<sup>®</sup> 70, 200-400 mesh, ammonium form) and lyophilised to give the title compound (222 mg, 86 %)

- 20  $[\alpha]_{\text{D}}^{22}$ : +55.7 $^\circ$  ( $c$  1.3, water)

- $^1\text{H}$  NMR data (pyridine- $d_5$ )  $\delta$ : 8.23-8.17 (2 H), 7.52-7.45 (1 H), 7.42-7.35 (2 H), 5.09 (dd, 1 H,  $J = 11.65$  and  $1.54$  Hz, H-6<sub>a</sub>), 4.82 (dd, 1 H,  $J = 11.43$  and  $6.81$  Hz, H-6<sub>b</sub>), 4.82 (s, 1 H, H-1), 4.66 (unresolved dd, 1 H,  $J \approx 3.5$  and  $3.5$  Hz, H-3),  
25 4.36 (dd, 1 H,  $J = 10.11$  and  $3.96$  Hz, H-4), 4.32-4.23 (1 H, H-5), 4.18 (d, 1 H,  $J = 3.08$  Hz, H-2), 3.60-3.44 (2 H,  $\text{CH}_2\text{CH}_3$ ), 3.16 (s, 3 H,  $\text{CH}_3\text{O}$ ) and 1.00 (t, 3 H,  $J = 7.0$  Hz,  $\text{CH}_2\text{CH}_3$ ) ppm.

- $^1\text{H}$  NMR data ( $\text{D}_2\text{O}$ , ref. acetone at 2.35 ppm)  $\delta$ : 8.22-8.17 (2  
30 H), 7.83 (tt, 1 H,  $J = 7.5$  and  $< 2$  Hz), 7.72-7.65 (2 H), 4.97 (HDO), 4.93 (s, 1 H, H-1), 4.81 (dd, virtually coupled, 1 H,  $J = 12$  and  $1.7$  Hz, H-6<sub>a</sub>), 4.65 (dd, virtually coupled, 1 H,  $J$

= 12 and 5.5 Hz, H-6<sub>b</sub>), 4.18-4.15 (2 H, h-4 and H-5), 4.05 (dd, 1 H, J = 3.3 and 1.4 Hz, H-2), 3.95-3.75 (sym. m. 2 H, CH<sub>2</sub>CH<sub>3</sub>), 3.94-3.91 (m, 1 H, H-3), 3.54 (s, 3 H, CH<sub>3</sub>O) and 1.35 (t, 3 H, J = 7 Hz, CH<sub>2</sub>CH<sub>3</sub>) ppm (the assignments of the <sup>1</sup>H NMR spectrum signals were made on the basis of a COSY experiment).

<sup>13</sup>C NMR data (D<sub>2</sub>O, ref.: acetone at 33.19 ppm) δ 171.4, 136.9, 132.5, 132.1, 131.7, 102.5, 78.4, 69.9, 69.1, 67.4, 65.7, 58.1, 56.0 and 17.5 ppm.

10 *Methyl 3-azido-6-O-pivaloyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl-α-D-altropyranoside*

A solution of methyl 3-azido-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl-α-D-altro-pyranoside (476 mg, 1.3 mmol) in pyridine (8 ml) was stirred at 0°C. Pivaloyl chloride (390 μl, 3.25 mmol) was added and the temperature was raised to room temperature. The stirring was continued for 1.5 h and the mixture was then cooled to 0°C. Methanol (10 ml) was added and the mixture was stirred for additional 20 min. at room temperature, evaporated and partitioned between ethyl acetate and water. The organic phase was washed subsequently with water, saturated aqueous NaHCO<sub>3</sub>, water, saturated aqueous NaCl, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and evaporated. Chromatography (SiO<sub>2</sub>, methyl tert-butyl ether-heptane 1:4) and (C-18 Lobar, Merck, acetonitrile-water 4:1) gave the pivaloate (431 mg, 75 %).

25 <sup>1</sup>H NMR data (CDCl<sub>3</sub>) δ: 7.28 and 6.89 (AB pattern, further coupled, 4 H, J = 8.6 Hz), 4.60 and 4.49 (AB pattern, 2 H, J = 11.0 Hz, ArCH<sub>2</sub>), 4.57 (s, 1H, H-1), 4.39 (dd, 1 H, J = 14.5 and 5.3 Hz, H-6<sub>a</sub>), 4.13 (2 H, H-6<sub>b</sub> and H-4), 3.93 (m, 1 H, H-3), 3.84-3.78 (4 H, H-5 and MeOAr), 3.64-3.45 (3 H, CH<sub>2</sub>CH<sub>3</sub> and H-2), 3.39 (s, 3H, 1-OMe) and 1.22-1.15 (s and t, 12 H, J = 7.0 Hz, C(CH<sub>3</sub>)<sub>3</sub> and CH<sub>2</sub>CH<sub>3</sub>) ppm.

<sup>13</sup>C NMR data (CDCl<sub>3</sub>) δ: 178.1, 159.6, 129.9, 129.3, 113.9, 99.5 (C-1), 76.7 (C-2), 72.5 (C-5), 71.9 (ArCH<sub>2</sub>), 66.4 (C-4),



123

66.2 ( $\text{CH}_2\text{CH}_3$ ), 63.6 (C-6), 58.5, 55.3, 55.2, 38.8, 27.2 and 15.4 ppm.

*Methyl 6-O-pivaloyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl-3-sulfamino- $\alpha$ -D-altropyranoside sodium salt*

- 5 (Cf. H. P. Wessel; *J. Carbohydr. Chem.* 11 (8), (1992), 1039-1052.)

To a stirred solution of methyl 3-azido-6-O-pivaloyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl- $\alpha$ -D-altropyranoside (300 mg, 0.683 mmol) in tetrahydrofuran (20 ml) was added water (150  
10 ml) and triphenylphosphine (0.90 g, 3.3 mmol). The solution was stirred for 24 h and was then concentrated to 5 ml. The residue was diluted with methanol (5 ml) and water (1 ml). One half of this solution was used to prepare the corresponding oxamate (see 48.182). To the remaining half of the  
15 stirred solution (0.342 mmol) was then added trimethylamine sulfurtrioxide complex (99 mg, 0.719 mmol). The pH was adjusted to 8.5 by 1M aqueous NaOH. After a few minutes a precipitate was formed, which was dissolved by adding tetrahydrofuran (0.7 ml) and methanol (1 ml). The reaction mixture  
20 was stirred for 0.5 h, after which time additional trimethylamine sulfurtrioxide complex (45 mg) was added. The stirring was continued for 15 min. During the course of the reaction pH was kept at 8-9 by adding 1M aqueous NaOH. The mixture was then diluted with water (2 ml) and the organic solvents were  
25 evaporated under reduced pressure. Additional water (5 ml) was added and the suspension was lyophilised and chromatographed ( $\text{SiO}_2$ , chloroform-methanol-water 100:15:1 $\rightarrow$ 80:20:1) to give the sulfamino compound, presumably in the sodium salt form, (161 mg, 80 %) after lyophilisation from water.

- 30  $^1\text{H}$  NMR data ( $\text{dmsO}-D_6$ )  $\delta$ : 7.24 and 6.87 (AB pattern, further coupled, 4 H,  $J = 8.8$  Hz), 4.69 and 4.24 (AB pattern, 2 H,  $J = 10.8$  Hz,  $\text{ArCH}_2$ ), 4.56 (s, 1H, H-1), 4.29 (dd, 1 H,  $J = 11.4$  and 1.7 Hz, H-6<sub>a</sub>), 4.05-3.96 (2 H, H-6<sub>b</sub> and NH), 3.83-3.76 (m, 2 H, H-2 and H-3), 3.76-3.73 (4 H, s and m, H-5 and

MeOAr), 3.54-3.40 (3 H,  $\text{CH}_2\text{CH}_3$  and H-4), 3.27 (s, 3H, 1-OMe) and 1.12-1.02 (s and t, 12 H,  $J = 7.0$  Hz,  $\text{C}(\text{CH}_3)_3$  and  $\text{CH}_2\text{CH}_3$ ) ppm.

5 *Methyl 6-O-pivaloyl-3-deoxy-2-O-ethyl-3-sulfamino- $\alpha$ -D-altro-pyranoside ammonium salt (bpy<sub>37</sub>)*

Methyl 6-O-pivaloyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl-3-sulfamino- $\alpha$ -D-altro-pyranoside sodium salt (153 mg, 0.259 mmol) was hydrogenolysed for 6 h in glacial acetic acid (8 ml) at 0.23 MPa and room temperature, using 5 % palladium on charcoal as a catalyst. Filtering and lyophilisation gave a residue which was chromatographed ( $\text{SiO}_2$ , chloroform-methanol-water 80:20:1 with 0.1%  $\text{NH}_3$ ) to give the title compound (77.2 mg, 74 %).

<sup>1</sup>H NMR data ( $\text{D}_2\text{O}$ , ref. acetone at 2.35 ppm)  $\delta$ : 4.83 (HDO), 15 4.78 (s 1 H, H-1), 4.43 (dd, virtually coupled, 1 H,  $J = 11.5$  and 1.1 Hz, H-6<sub>a</sub>), 4.30 (dd, virtually coupled, 1 H,  $J = 11.5$  and 5.1 Hz, H-6<sub>b</sub>), 3.95-3.91 (2 H, H-4 and H-5), 3.89 (dd, 1 H,  $J = 3.3$  and 1.3 Hz, H-2), 3.80-3.62 (m. 3 H, H-3 and  $\text{CH}_2\text{CH}_3$ ), 3.42 (s, 3 H,  $\text{CH}_3\text{O}$ ) and 1.24-1.21 (s and t, 12 H,  $J = 7$  Hz, <sup>t</sup>Bu and  $\text{CH}_2\text{CH}_3$ ) ppm (the assignments of the <sup>1</sup>H NMR 20 spectrum signals were made on the basis of a COSY experiment).

<sup>13</sup>C NMR data ( $\text{D}_2\text{O}$ , ref.: acetone at 33.19 ppm)  $\delta$  188.3, 106.2, 82.2, 73.8, 72.8, 70.6, 69.3, 61.9, 59.8, 45.6, 37.1 and 21.3 25 ppm.

*Methyl 6-O-pivaloyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl-3-<sup>t</sup>butyloxamido- $\alpha$ -D-altropyranoside*

Oxalyl chloride (120  $\mu\text{l}$ , 1.37 mmol) dissolved in tetrahydrofuran (1 ml) was cooled to -26°C. A solution of *t*-butanol 30 (131 mg, 1.77 mmol) and pyridine (154  $\mu\text{l}$ , 1.91 mmol) in tetrahydrofuran (2 ml) was added dropwise and the mixture

was stirred at -26°C for 15 min, while the initially formed yellowish precipitate turned white.

The solvents were removed under reduced pressure from the tetrahydrofuran solution from the reduction of 3-azido-6-O-pivaloyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl- $\alpha$ -D-altropyranoside (150 mg, 0.342 mmol). The residue was redissolved in tetrahydrofuran (2.5 ml) and pyridine (154  $\mu$ l) and added dropwise to the cooled solution above and stirred at -26°C for 1h. Water (2 ml) was added and the mixture diluted with t-butyl methyl ether (15 ml) and washed with water, saturated sodium hydrogen carbonate and brine. The organic solution was dried over sodium sulfate and the solvents were removed under reduced pressure. The residue was chromatographed (SiO<sub>2</sub>, t-butyl methyl ether - toluene 1:8) to yield 187 mg, 100% of the title compound.

<sup>1</sup>H NMR data (CDCl<sub>3</sub>)  $\delta$ : 8.25 (d, 1H, 10.1 Hz, 3-NH), 7.26 and 6.85 (AB pattern, further coupled, 4 H, J = 8.6 Hz), 4.82-4.71 (m and d, 2H, 10.5 Hz, H-3 and ArCH<sub>2</sub>), 4.69 (s, 1H, H-1), 4.42 (dd, 1H, 1.7 and 11.7 Hz, H-6<sub>a</sub>), 4.31 (d, 1H, 10.5 Hz, ArCH<sub>2</sub>), 4.16 (dd, 1H, 6.2 and 11.7 Hz, H-6<sub>b</sub>), 3.88 (ddd, 1H, 1.8, 5.9 and 10.3 Hz, H-5), 3.81-3.75 (dd and s, 4H, H-4 and ArOMe), 3.64-3.49 (m, 2H, CH<sub>2</sub>CH<sub>3</sub>), 3.47-3.43 (dd and s, 4H, 1.3 and 3.1 Hz, H-2 and 1-OMe), 1.55 (s, 9H, oxamate C(CH<sub>3</sub>)<sub>3</sub>), 1.19 and 1.18 (s and t, 12H, 7 Hz, 6-O pivaloate <sup>t</sup>Bu and CH<sub>2</sub>CH<sub>3</sub>) ppm.

<sup>13</sup>C NMR data (CDCl<sub>3</sub>)  $\delta$ : 178.1, 159.4, 159.1, 157.4, 130.3, 129.4, 113.8, 98.9, 84.2, 77.2, 75.9, 70.8, 69.5, 65.8, 65.7, 63.4, 55.2, 55.1, 45.7, 38.8, 27.7, 27.1, 15.3 ppm.

Methyl 6-O-pivaloyl-3-deoxy-2-O-ethyl-3-oxamido- $\alpha$ -D-altropyranoside ammonium salt (bpy<sub>54</sub>)

Methyl 6-O-pivaloyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl-3-<sup>t</sup>butyloxamido- $\alpha$ -D-altropyranoside (180 mg, 332  $\mu$ mol) was dissolved in acetic acid-ethanol mixture (1:1, 6 ml), palladium on carbon (5%, 250 mg) was added and subjected for

hydrogenolysis at 39 psi overnight. The catalyst was filtered of on a celite pad, washed with ethanol. The filtrate was concentrated under reduced pressure, redissolved in dichloromethane (2.4 ml). Trifluoroacetic acid (960  $\mu$ l) was added and the mixture stirred for 4.5h at room temperature. Water (3 ml) and ammonia (3M, 300  $\mu$ l) was added and reaction mixture was concentrated under reduced pressure and lyophilized. The residue was chromatographed ( $\text{SiO}_2$ , chloroform-methanol-water 80:20:1 with 0.1%  $\text{NH}_3$ ) to give the title compound, 117.1 mg, 89%.

$^1\text{H}$  NMR data ( $\text{D}_2\text{O}$ , ref. acetone at 2.35 ppm)  $\delta$ : 4.80 (HDO), 4.78 (s 1 H, H-1), 4.45-4.35 (m, 2H, H-3 and H-6<sub>a</sub>), 4.26 (dd, 1H, 4.4 and 12.1 Hz, H-6<sub>b</sub>), 4.07-3.96 (m, 2H, H-4 and H-5), 3.77-3.58 (m, 3H,  $\text{CH}_2\text{CH}_3$  and H-2), 3.45 (s, 3H, 1-OMe), 1.20-1.13 (s and t, 12 H, 6-O-pivaloate  $^t\text{Bu}$  and  $\text{CH}_2\text{CH}_3$ ) ppm.

$^{13}\text{C}$  NMR data ( $\text{D}_2\text{O}$ , ref.: acetone at 33.19 ppm)  $\delta$ : 188.1, 171.9, 171.4, 105.7, 82.4, 73.7, 72.9, 70.3, 69.7, 62.0, 56.1, 45.9, 37.0, 21.2 ppm.

*Methyl 3-azido-6-O-pyrrol-3'-ylcarboxyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl- $\alpha$ -D-altropyranoside*

A solution of methyl 3-azido-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl- $\alpha$ -D-altro-pyranoside (476 mg, 1.3 mmol) in pyridine (496  $\mu$ l) and dichloromethane (2 ml) was stirred at 0°C. Methanesulfonyl chloride (380  $\mu$ l, 4.90 mmol) was added and the temperature was raised to room temperature. The stirring was continued for 5 h and the mixture was then cooled to 0°C. Water (1 ml) was added and the mixture was stirred for additional 20 min. at room temperature, evaporated and partitioned between ethyl acetate and water. The organic phase was washed subsequently with water, saturated aqueous  $\text{NaHCO}_3$ , water, saturated aqueous  $\text{NaCl}$ , dried ( $\text{Na}_2\text{SO}_4$ ), filtered and evaporated. Chromatography ( $\text{SiO}_2$ , methyl tert-butyl ether-toluene 1:9) yielded 216 mg (79%) of the mesylate as a yellowish oil.

N-Triisopropylsilyl pyrrole-3-carboxylic acid (484 mg, 1.81 mmol) and 1,8-diazacyclo[5.4.0]undec-7-ene (255  $\mu$ l) were dissolved in dimethylformamide (1 ml) and added to a solution of the mesylate above in dimethylformamide (1 ml).

- 5 The mixture was heated at 90°C overnight and the subjected to chromatography (210 g C-8 Lobar, Merck, acetonitrile-water 3:2) to yield 30 mg, 16 % of the pyrroloyl ester.

<sup>1</sup>H NMR data (CDCl<sub>3</sub>)  $\delta$ : 8.73 (br.s., 1H, pyrrole NH), 7.39 (symm. m., 1H, pyrrole H-4), 7.27 and 6.85 (AB pattern, 10 further coupled, 4 H, J = 8.8 Hz), 3.74 (symm. m, 1H, pyrrole H-2), 3.64 (symm. m, 1H, pyrrole H-5), 4.63-4.49 (s and 2 d, 3H, H-1 and ArCH<sub>2</sub>), 4.48-4.36 (m, 2H, 2 H-6), 4.22 (symm. m, 1H, H-5), 3.94-3.84 (m, 2H, H-3 and H-4), 3.78 (s, 3H, 15 ArOMe), 3.68-3.55 (m, 3H, CH<sub>2</sub>CH<sub>3</sub> and H-2), 3.40 (s, 3H, 1-OMe), 1.20 (t, 3H, 6.8 Hz, CH<sub>2</sub>CH<sub>3</sub>) ppm.

<sup>13</sup>C NMR data (CDCl<sub>3</sub>)  $\delta$ : 164.6, 159.5, 129.9, 129.4, 123.6, 11.7, 116.1, 113.9, 109.9, 99.9, 76.9, 72.5, 72.0, 66.8, 66.3, 63.0, 58.9, 55.3, 55.2, 15.4 ppm.

20 Methyl 6-O-pyrrol-3'-ylcarboxyl-3-deoxy-2-O-ethyl-3-sulfamino- $\alpha$ -D-altropyranoside ammonium salt (bpy<sub>40</sub>)

To a stirred solution of methyl 3-azido-6-O-pivaloyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl- $\alpha$ -D-altropyranoside (30 mg, 61  $\mu$ mol) in tetrahydrofuran (1 ml) was added water (11  $\mu$ l) and triphenylphosphine (83 mg, 305  $\mu$ mol). The solution was 25 stirred for 24 h, diluted with and water (1 ml). Trimethylamine sulfurtrioxide complex (9 mg, 0.72  $\mu$ mol) was then added. The pH was adjusted to 8.5 by 1 M aqueous NaOH. After a few minutes a precipitate was formed, which was dissolved by adding tetrahydrofuran (0.7 ml) and methanol (1 ml). The 30 reaction mixture was stirred for 0.5 h, after which time additional trimethylamine sulfurtrioxide complex (9 mg) was added. The stirring was continued for 15 min. During the course of the reaction pH was kept at 8-9 by adding 1M aqueous NaOH. The mixture was then diluted with water (2 ml),

washed with t-butyl methyl ether, lyophilised and chromatographed ( $\text{SiO}_2$ , chloroform-methanol-water 80:20:1) to give the sulfamino compound, presumably in the sodium salt form, (30.6 mg, 93 %) after lyophilisation from water.

- 5  $^1\text{H}$  NMR data ( $\text{CD}_3\text{OD}$ )  $\delta$ : 7.35 (unresol. dd, x part of a ABX-pattern, 1H,  $J_{\text{AX}}+J_{\text{BX}}=3.5$  Hz, pyrrole H-2), 7.29 and 6.79 (AB pattern, further coupled, 4 H,  $J = 8.6$  Hz), 6.75 (dd, 1H, 2.0 and 2.9 Hz, pyrrole H-4), 6.51 (dd, 1H, 1.3 and 2.9 Hz, pyrrole H-5), 4.71 and 4.46 (AB pattern, 2 H,  $J = 11.4$  Hz, ArCH<sub>2</sub>), 4.66 (bs, 1 H, H-1), 4.39 and 4.33 (2 dd, 2H, 2.6, 5.1 and 11.7 Hz, 2 H-6), 4.04 (m, 1H, H-3), 3.97-3.87 (m and dd, 2H, 1.1 and 3.1 Hz, H-5 and H-2), 3.80 (dd, 1H, 4.4 and 10.1 Hz, H-4) 3.73 (s, 3H, ArOMe), 3.68-3.52 (m, 2H, CH<sub>2</sub>CH<sub>3</sub>), 3.39 (s, 3H, 1-OMe), 1.15 (t, 3H, 7.0Hz, CH<sub>2</sub>CH<sub>3</sub>) ppm.
- 10
- 15  $^{13}\text{C}$  NMR data ( $\text{CD}_3\text{OD}$ )  $\delta$ : 167.32, 160.9, 131.3, 131.0 125.3, 120.1, 116.3, 114.7, 110.2, 101.7, 77.2, 10.4, 69.9, 67.2, 66.4, 55.6, 55.5, 50.9 and 15.8 ppm.

- The residue was dissolved in methanol (15 ml), palladium on carbon (5%, 50 mg) was added and the mixture hydrogenolyzed at 30 psi for 3 h. The catalyst was filtered off, the filtrate concentrated under reduced pressure, redissolved in ammonia (2%) in methanol-water (1:1, 0.6 ml) and chromatographed (3 ml C-18 Supelco, methanol-water 2:3). Methanol was removed under reduced pressure from the pooled product-containing fractions and the residue lyophilized to give 6.2 mg of the title compound as the ammonium salt.
- 20
- 25

- $^1\text{H}$  NMR data ( $\text{CD}_3\text{OD}$ )  $\delta$ : 7.35 (m, 1H, pyrrole H-2), 6.75 (dd, 1H, 2.0 and 2.9 Hz, pyrrole H-4), 6.55 (dd, 1H, 1.3 and 2.9 Hz, pyrrole H-5), 4.67 (br.s, 1H, H1), 4.55 (dd, 1H, 1.3 and 11.4 Hz, H-6<sub>a</sub>), 4.31 (dd, 1H, 5.9 and 11.6 Hz, H-6<sub>b</sub>), 3.86-3.76 (m, 4H, H-2, H-3, H-4, H-5), 3.75-3.56 (m, 2H, CH<sub>2</sub>CH<sub>3</sub>), 3.40 (s, 3H, 1-OMe), 1.21 (t, 3H, Hz, CH<sub>2</sub>CH<sub>3</sub>) ppm.
- 30

## EXAMPLE 6

*Chaperone assays and the use thereof in testing compounds for inhibition of chaperone binding*

In order to conclusively identify peptides or peptide mimetics that bind strongly to *PapD*, a good assay is required. The maltose binding protein (MBP) has been engineered by the inventors so that it is recognized by *PapD*. The commercially available plasmid pMAL-P2 encoding MBP with a linker arm encoding sequence fused to the 3' end of the gene was used. MBP is secreted into the periplasmic space and is easily purified by affinity chromatography using amylose resin. MBP is eluted from the column with 20 mM maltose. *PapD* is not co-eluted with MBP when co-expressed with MBP whereas fusing the *PapG* COOH-terminus to MBP results in *PapD* binding to MBP to form a complex that is co-eluted from the affinity column. Using this strategy, it has been found that fusing the carboxy terminal 140 (MBP-G1'-140') and 134 (MBP-G134') amino acids of *PapG* to MBP results in the formation of a strong *PapD*-MBP complex which is purified by maltose affinity chromatography. The complex between the MBP fusion and *PapD* has been shown to be stable in up to 3 M urea similar to the stability of a *PapD*-*PapG* complex arguing that the COOH-terminal 134 amino acids of *PapG* probably contains most of the chaperone recognition motif. In addition, binding to the MBP fusion protein has proven dependent on the universal chaperone anchoring residues, Arg-8 and Lys-112, as described herein since mutations in these two residues abolished *PapD* binding to the MBP fusion protein.

It is also interesting to test the ability of MBP-G fusions to inhibit pilus assembly and thus bacterial attachment by co-expressing the fusion proteins in cells producing P pili. If *PapD* binds to the MBP-G fusion, it will be titrated away from pilus subunits and thus decrease or abolish pilus formation. This technique will be one mechanism to validate the concept that chaperone binding peptides can abolish the

virulence of a pathogen by preventing the assembly of surface-localized adhesins. Using this analysis it has been found that in strains producing P pili, co-expression of the MBP-G1'-140' and MBP-G134' fusion completely abolishes the ability of the bacteria to bind red blood cells and cause hemagglutination. Further, it has recently been shown by the inventors that after coexpression of the MBP-G1'-140' fusion in HB101/pPAP5-bacteria, it was not possible to detect any visual signs of pili on the surface of the bacteria by electron microscopy. This is evidence for the theory expressed in this invention, that the prevention/inhibition of binding between pilus subunits and molecular chaperones also prevents the assembly of intact pili.

These results support the biological relevance of the crystal structure described in Fig. 3, showing that other proteins can be engineered to be recognized by *PapD* by fusing the carboxy terminal recognition site onto that protein. The MBP-G134' and MBP-G1'-140' fusions have been purified to establish *in vitro* assays to measure interactions with *PapD*. The purified fusion proteins are coated on wells of microtiter plates and the ability of *PapD* to bind to the fusion proteins is tested in an ELISA procedure. This assay is critical since it allows testing the ability of compounds to inhibit binding of *PapD* to the domain on *PapG* recognized by the chaperone. Furthermore, the purified fusion proteins can be used in the Pharmacia BiaCore® assay to quantitate binding to *PapD* and establish inhibition assays for compounds of the invention. Carboxyl-terminal peptides have already been suggested to correspond to part of the chaperone recognition motif. One can thus investigate whether the carboxyl terminal peptides are able to inhibit binding of *PapD* to the MBP-G1'-140' fusion in the ELISA assay. The development of this high through-put assay makes it possible to screen peptide libraries, chemical libraries, natural compounds, and peptide mimetics for their ability to inhibit chaperone binding. In addition, this assay can be used to test whether the known periplasmic chaperones utilize common recognition paradigms.



In addition to testing compounds for their ability to inhibit binding to the MBP-G fusion, other assays have been established which measure PapD-peptide interactions. For example, an ELISA has been developed to measure the binding of PapD to peptides coated on wells of microtiter plates. In this assay the octameric peptide G1'-8'WT was an equally efficient inhibitor as the 19-mer G1'-19'WT, revealing the octamer to be the optimal starting point for designing of modified compounds with increased affinity for the binding site of PapD (cf. Example 7).

Also, as previously described, native PapD is able to bind to reduced, denatured PapG and restore the PapD-PapG complex in vitro in a reconstitution assay (Kuehn et al., 1991). This assay has been proposed to reflect the recognition function of PapD in vivo and will be used to determine the ability of the compounds to inhibit PapD binding to PapG in vitro. For example, the carboxyl terminal PapG peptide has been shown to inhibit the binding of PapD to PapG in this assay, establishing that it occupies the pilus subunit binding site of PapD. Another way to map PapD-peptide interactions is to test the ability of compounds to reduce the rate of well defined proteolytic cleavage events. For example, partial digestion with trypsin cleaves PapD in the F1-G1 loop at residue Lys-99 (See "T" site denoted in Fig. 3). The rate of tryptic cleavage of PapD was reduced by preincubation of PapD with PapG and PapK peptides (cf. Example 2). The observed protection of PapD by bound peptides may be due to a change of the local conformation of the F1-G1 loop, or due to physical contact of the loop by the peptide. These assays may thus be used as an initial screening for the ability of new compounds to bind to PapD and interfere with its recognition function. Strong binding compounds that inhibit PapD binding to the MBP-G1'-140' fusion will be co-crystallized with PapD to provide the structural basis of the recognition surface used by PapD. As new crystallographic data becomes available, the relevance of critical PapD-inhibitor or PapD-enhancer interactions will be tested by determining the effect of site-directed muta-

tions on the ability of the chaperone to bind pilus subunits and mediate pilus assembly. This important information will lead to defining how chaperones recognize subunits and the function of chaperone-subunit interactions.

5 A quantitative chaperone binding assay has also been developed: A modified *PapG* peptide has been synthesized, wherein Ser-9' is substituted by a Cys. From the crystal structure of *PapD* bound to the peptide, the side chain of Ser-9' was not predicted to interact with *PapD*. Instead, this side chain was  
10 oriented towards the solvent, but Ser-9' was adjacent to the last amino acid of the interactive zipper between *PapD* and the peptide. The environmentally sensitive fluorescent probe 5-IAF (5-iodoacetamidfluorescein) was then covalently coupled to the peptide via the sulfhydryl group on the Cys residue  
15 and the labelled peptide was purified; of course any other suitable environmentally sensitive fluorescent probe may be employed. It was found that the addition of *PapD* to the peptide causes a marked decrease in the fluorescent intensity and a shift in the emission maxima. This information was used  
20 to calculate a binding constant for the *PapD*-peptide interaction. The change in the fluorescence intensity at 514 nm after the addition of increasing concentrations of *PapD* was determined. By plotting the fluorescent intensities versus *PapD* concentrations, a binding constant of  $2.5 \times 10^{-6}$  M was  
25 calculated (The calculations were performed on the commercially available computer program "Kaleidagraph").

It is thus possible to evaluate the binding constants of other substances binding to *PapD* in a quantitative manner, as the addition of a substance which binds competitively to *PapD*  
30 will result in an increased fluorescence compared to a situation where less or no competitive substance is present in the system.

A similar assay system will be developed using MBP-G1'-140' instead of the *PapG* peptide as described herein. As this  
35 fusion protein has been found to interact with two binding

sites in *PapD*, it will be possible to 1) quantitate the binding affinity to the second binding site (on domain 2) and 2) screen for compounds which interact with either of the two binding sites and quantitate the interaction. Of course, 5 labelling the site 2 peptide with a suitable environmentally sensitive probe should allow a determination of the binding constant using the same methodology as describe above.

The discovery that the DegP protease is greatly responsible for the degradation of pilin subunits in the absence of a 10 chaperone renders the *degP41* strain an interesting candidate for an *in vivo* test system of the effects exerted on *PapD* by compounds of the invention.

When administering a substance which prevents, inhibits or enhances the binding between the chaperone and the pilus 15 subunits to a system containing the *degP41* strain, the substance should, even in small amounts, be toxic to the bacteria, as the DegP<sup>-</sup> bacteria will be incapable of degrading the accumulating pilus subunits. Note that a *degP41* strain was used to identify a hitherto unknown binding site 20 on domain 2 of *PapD*, cf. example 10.

However, such an assay requires that the tested substance is able to enter the periplasmic space before it can exert its effect; this fact renders this type of assay less suitable as a screening assay for lead compounds, as it will "ignore" 25 substances which have the desired effects on chaperone-subunit interaction, but which are e.g. too hydrophillic to enter the periplasmic space. On the other hand, the system will be well-suited for assessing the clinical potential of substances which have already proven successful in the *in* 30 *vitro* assays described herein.

Once the model *PapD* system is in place, the experiments described above can of course be expanded to include the other members of the *PapD*-like family of chaperones. In this way it will be possible to establish the general requirements

for chaperone recognition and the molecular basis of the chaperone recognition paradigm in gram negative bacteria.

Some specific assays for determination of interactions with molecular chaperones are described in detail in Example 10.

## 5 EXAMPLE 7

*Design of peptides and peptide mimetics capable of binding to the binding site of PapD and other pilus assembling chaperones*

To provide further information on chaperone-subunit interactions, it is contemplated to synthesize peptides overlapping the entire PapG protein and other pilus subunits. In addition, probing peptide binding to available chaperones using both a chemical peptide library and a phage display library should be performed. The two libraries are complementary since the peptides are presented in different environments which could importantly influence binding to the chaperones. In practice the chemical library is limited to all possible hexapeptides and to hepta- and octapeptides if some of the natural amino acids are excluded or if one or two residues are kept invariant. This library can be evaluated by direct binding of the chaperone to the resin beads and sequencing of interesting peptides. The phage display library contains approximately  $10^{10}$  peptides in the PIII protein of the so-called "fusion phage" that retains phage function and displays the foreign peptides on the surface. The ability of peptide-containing phages to bind PapD coated in wells of microtiter plates can be detected in an ELISA using the techniques described in Example 6. Positive binding phages will be amplified, purified and retested for their ability to bind PapD. The sequence of the peptide insert of positively binding phages will be determined and the corresponding peptides will be synthesized and tested for their ability to inhibit binding of PapD to the MBP-G fusion as described in Example 6.

The results from the studies outlined above will be of the utmost importance for design and evaluation of ligands acting as "chaperone inhibitors". In spite of this, the crystal structure of the *PapD*-*PapG* 19-mer peptide complex provides  
5 sufficient insight into chaperone-ligand interactions to initiate systematic studies of small peptides and peptide mimetics as chaperone inhibitors already at the present stage. The high inhibitory power of the *PapG* 8-mer peptide to *PapD* (equal to the *PapG* 19-mer peptide, cf. Example 6)  
10 reveals the feasibility of such studies and the conserved features of the proposed binding site on the chaperones indicate that such inhibitors could well have broad specificity.

The specificity in recognition of the *PapG* 19-mer by *PapD* has  
15 been suggested to be provided by anchoring of the peptide's carboxyl terminus to the *PapD* cleft residues Arg-8 and Lys-112 and subsequent "zippering" interactions between alternating hydrophobic residues in the peptide and complementary hydrophobic residues in *PapD*. The following experiments  
20 support this hypothesis. Firstly, mutations in the anchor residues Arg-8 and Lys-112 have abolished subunit binding *in vivo*. Secondly, deletion of the C-terminal residue in the *PapG* 19-mer leads to a substantially decreased binding to *PapD* as the hydrophobic "zippering" interactions between  
25 peptide and *PapD* were then placed off register when the C-terminus of the deletion peptide was anchored to Arg-8 and Lys-112. The "zippering" hypothesis for specificity in binding between *PapD* and peptides (or pilus subunits) has been preliminary investigated by performing the following experiments:  
30

A length series of peptides from the C-terminus of *PapG* consisting of G1'-5', G1'-6', G1'-7', G1'-8', G1'-11', G1'-16' and G1'-19' was prepared. A replacement series and a deletion series of G1'-8' was also synthesized. In the replacement series Pro-1', Phe-2', Leu-4', Val-5', Met-6' and  
35 Met-8' was replaced by Ser, whereas Ser-3' and Thr-7' was

replaced by Ala (i.e. hydrophobic amino acids was replaced by the polar Ser and hydrophilic amino acids was replaced by Ala). In the deletion series one residue at a time in G1'-8' was deleted with simultaneous addition of a serine (which is found in position 9' in native PapG) to maintain the peptide length at 8 amino acids. The peptides were all synthesized by the Fmoc solid phase strategy and were purified by reversed phase HPLC.

The ability of the peptides to inhibit the binding of PapD to the MBP-G1'-140' fusion protein was then investigated using the following ELISA test:

Stock solutions of MBP-G1'-140' proteins in PBS were diluted to 0.1  $\mu$ M with PBS. The wells of microtiter plates were coated with 50  $\mu$ l of the protein solutions overnight at 4°C. The wells were washed with PBS, and blocked with 200  $\mu$ l of 3% bovine serum albumin (BSA) in PBS for 2 hours at 25°C. The plates were washed vigorously three times with PBS and incubated with 50  $\mu$ l of 1-5  $\mu$ M PapD proteins in 3% BSA-PBS for 45 minutes at 25°C. The PapD was preincubated with each of the peptides at a 1:25 ratio for 30 min before being added to the wells. After three washings with PBS, the wells were incubated with a 1:500 dilution of rabbit anti-PapD antiserum in 3% BSA-PBS for 45 minutes at 25°C. After three washings with PBS, the wells were incubated with a 1:1000 dilution of goat antiserum to rabbit IgG coupled to alkaline phosphatase in 3% BSA-PBS for 45 minutes at 25 °C. After three washings with PBS and three washings with developing buffer (10 mM diethanolamine, 0.5 mM  $MgCl_2$ ), 50  $\mu$ l filtered 1 mg/ml p-nitrophenyl phosphate in developing buffer was added, the reaction was incubated for 1 hour or longer if necessary in the dark at 25°C, and the absorbance at 405 nm was read.

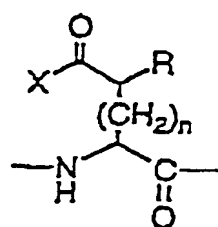
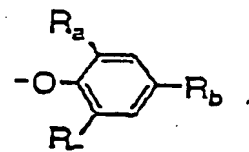
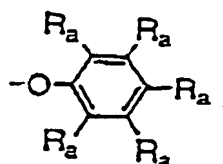
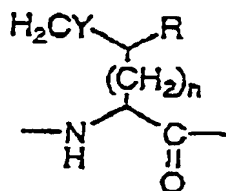
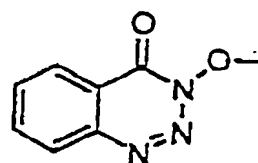
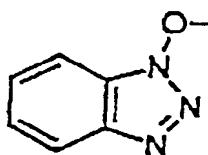
The inhibitory powers of the peptides in the three series are presented in Figs. 17-19 and the number of experiments performed with each series is given in the Figs. Vertical lines for each peptide in the figures are 95% confidence intervals

obtained after a statistical analysis of the experimental data. As revealed by the evaluation of the length series, the peptides G1'-8', G1'-11' and G1'-16' are significantly more potent than the shorter G1'-6' and G1'-7' (Fig. 17). This observation fits very well with the crystal structure of PapD complexed with G1'-19' which shows that the C-terminal 8 residues in G1'-19' are hydrogen bonded to PapD. The shorter peptides G1'-6' and G1'-7' are unable to fulfil this hydrogen bonding pattern and are therefore less active inhibitory peptides (Fig. 18). The replacement series reveals that residues 4', 5' and 6' in G1'-8' form important contacts with PapD, since their replacement results in less active inhibitory peptides (Fig. 18). The deletion series again indicated an important role for residues 4', 5' and 6' in G1'-8' for the complex formation with PapD (Fig. 19). However, the results obtained with the deletion series did not support the "zippering" hypothesis according to which the members of the deletion series would show an increasing inhibitory potency as the deletion is moved from the C-terminus towards the N-terminus of G1'-8'.

As appears from these preliminary results, the results obtained in the assay exhibit large deviations from the mean in each experiment. As discussed herein, one reason for this might be the slow kinetics of binding between PapD and correctly folded pilus subunit proteins and analogues of such correctly folded subunit proteins. It is therefore contemplated to modify the assay by introducing denaturing influences powerful enough to at least partially unfold the pilus subunit (analogues). It is expected that this will reduce the deviations in the assay results. Another reason for the large deviations might be binding of the peptides to BSA used in the ELISA. Therefore, the replacement of BSA with other macromolecules will be investigated.

- In case the "zipper mechanism" for peptide binding to *PapD* can be confirmed, peptide chaperone binding can be optimized using a limited synthetic peptide library in which Pro-1', Phe-2', Leu-4', Met-6' and Met-8' of the *PapG* 8-mer are
- 5 replaced by the hydrophobic amino acids Val, Leu, Ile, Met, Phe, Trp, Tyr and His. Non-natural amino acids such as D-amino acids and N-methylated amino acids, as well as amino acids containing aliphatic and various aromatic side chains, will also be incorporated in the chemical library. Optimal
- 10 amino acid combinations for these five positions will be used in the synthesis of the chaperone inhibitors described below in this Example. If the "zipper mechanism" is not confirmed, the outlined approaches will be applied to residues shown to be important for binding to *PapD*.
- 15 Chaperone inhibitors that form covalent bonds to the chaperone after docking into its active site will be developed. The crystal structure of the *PapD*-peptide complex shows that the C-terminal carboxyl group of the peptide is hydrogen bonded to Lys-112 and Arg-8 in the chaperone and that the side chain
- 20 of Val-5' in the peptide is close to the side chain amino group of another Lys in *PapD*. Introduction of reactive groups such as alkyl halides, aldehydes, acid halides and active esters in these positions of the optimized 8-mer peptide will lead to the formation of covalent bonds to the lysines in
- 25 *PapD*, and the peptide derivatives thus constitute high affinity inhibitors. The inhibitors are based on nonamer or shorter peptides in which residue 5 (from the C-terminus) and/or the C-terminal carboxyl group have been modified as indicated:

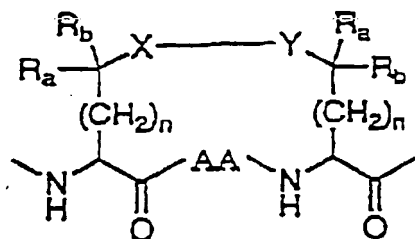


*Replacements for residue 5 from the C-terminus:* $n = 0-5$  $R = H, Me, Et, C_3H_7, C_4H_9, C_6H_5$  $X = H, F, Cl, Br,$  $R_a = F, Cl, Br$  $R_a = R_c = H, R_b = NO_2$  $R_a = R_b = NO_2, R_c = H$  $R_a = R_b = R_c = NO_2$  $n$  and  $R$  as above $Y = Cl, Br, I, OSO_2Me, OSO_2CF_3, OSO_2C_6H_4Me$ *Replacements for the C-terminal COOH-group:* $-COX$ ,  $X$  as above $-CH_2Y$ ,  $Y$  as above

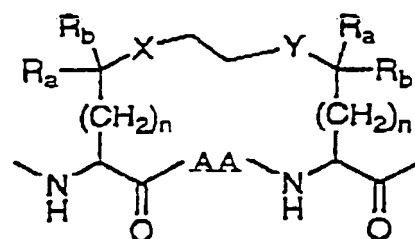
Such interactions between aldehydes and lysine side chains have given potent drug candidates for sickle-cell anaemia.

In the PapD-peptide crystal, the peptide forms an extension of a  $\beta$ -sheet in the chaperone. Restrictions that give the peptide a  $\beta$ -sheet like conformation will therefore result in a favourable change in the entropy of binding. Conformation-

ally restricted peptides constituting miniature cyclic  $\beta$ -sheets or peptides having the side chains of amino acids separated by one amino acid covalently linked to each other will be prepared based on the optimized 8-mer peptide. Peptides with covalently linked side chains will have the three consecutive amino acids replaced by the fragments shown below:



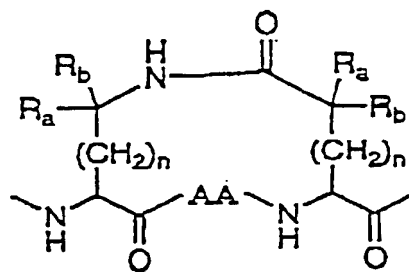
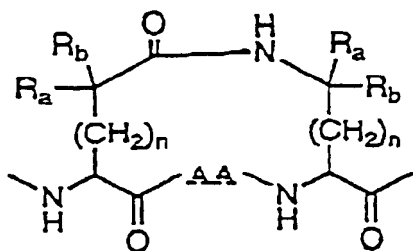
$X, Y = SS; S, CH_2; CH_2S; O, CH_2; CH_2, O; CH_2CH_2$



$X, Y = SS; S, CH_2; CH_2S; O, O; O, CH_2; CH_2, O; CH_2, CH_2$

AA = any natural or nonnatural amino acid

$R_a, R_b$  = all possible combinations of H, Me, Et,  $C_3H_7$ ,  $C_4H_9$ ,  $C_6H_5$

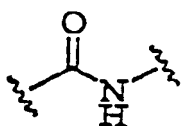


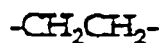
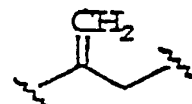
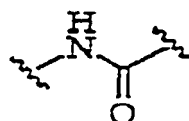
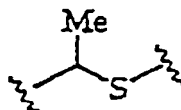
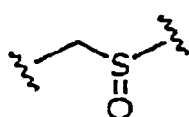
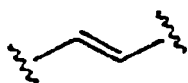
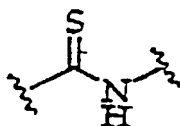
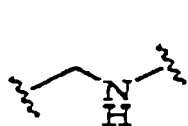
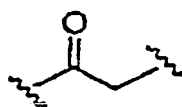
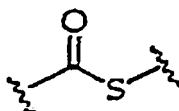
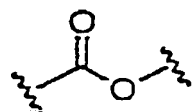
AA = any natural or nonnatural amino acid

$R_a, R_b$  = all possible combinations of H, Me, Et,  $C_3H_7$ ,  $C_4H_9$ ,  $C_6H_5$

Interesting inhibitors will be co-crystallized with chaperones, and the complexes will also be investigated by NMR spectroscopy.

Peptides and peptide mimetics used as drugs can be rapidly  
5 metabolized by proteolytic enzymes that cleave the peptide  
bonds. Chymotrypsin selectively cleaves the peptide bond on  
the carboxyl side of amino acids with aromatic and large  
hydrophobic side chains. In the PapG 8-mer, the Met-8'-  
Thr-7', Met-6'-Val-5' and Phe-2'-Pro-1' amide bonds are  
10 therefore especially sensitive to proteolysis and should be  
replaced by metabolically stable peptide isosters. Examples  
of such peptide isosters are given below:

Replacements for the  fragment:



#### EXAMPLE 8

The role of periplasmic chaperones in the infectivity of bacteria which adhere by the means of pili

The specific role of chaperone-assisted pilus assembly in  
 5 virulence can be determined by comparing the adherence and  
 pathogenicity of pilated wild type and isogenic mutants that  
 are non-piliated due to site-directed mutations in the active  
 site of the chaperone. Mutations in residues such as Arg-8  
 and Lys-112 that constitute the subunit binding site of *PapD*  
 10 can be recombined into the bacterial chromosome of the clini-

cal isolate strain DS17 using the same method that has been successful to introduce point mutations into the papG gene in the chromosome of the same strain. *E. coli* strain DS17 spread epidemically within a neonatal ward, causing several cases of pyelonephritis. In addition, the strain has been found to cause acute kidney infections in a cynomolgus monkey pyelonephritis model. DS17 contains one pap gene cluster with a papG gene expressing a typical papG adhesin. It also expresses typical type 1 pili containing the mannose-binding adhesin, FimH. Type 1 pili have been suggested to be important virulence determinants in cystitis. Therefore, site-directed mutants in the subunit binding site of the *FimC* chaperone should also be generated to test this concept.

It is contemplated that a compound that binds the chaperone active site would block pilus assembly and thus prevent attachment. If this concept is valid, then well-defined chaperone mutants such as in Arg-8 and Lys-112 should abolish receptor binding, too. The receptor binding activity of various chaperone mutants should be measured in the receptor binding ELISA assays described in examples 2 and 6. To measure type 1 pilus-mediated binding, mannose will be linked to the wells of the microtiter plates. Adherence of various *FimC* and *PapD* chaperone mutants to the immobilized receptors will be quantitated using antibodies to *E. coli* DS17 in an ELISA experiment. Furthermore, the ability of small *PapD*-binding peptides to block pilus assembly in DS17 and, thus, to abolish receptor binding will be tested. Strain DS17 will be grown in the presence of short *PapD*-binding or non-binding peptides and then tested for their ability to bind receptor in the ELISA experiment. These experiments will make important strides at validation the concept that anti-chaperone inhibitors would prevent bacterial attachment *in vitro*.

Then, the role of chaperone-assembled adhesins in causing diseases will be established. The causality between *E. coli* expressing P pili and pyelonephritis has until recently been based solely on epidemiological data. It has been shown that

the P piliated strain DS17 causes pyelonephritis to occur in the normal urinary tract of cynomolgus monkeys. A mutation in the papG gene of strain DS17 was generated by allelic replacement to create the derivative strain DS17-8 to test the requirement of the PapG adhesin in causing the disease. DS17-8 contains one base pair deletion after codon 37 in papG resulting in the expression of P-pili lacking the PapG tip adhesin. This mutant was unable to bind the globoside receptor *in vitro* and was unable to bind to kidney tissue from humans or cynomolgus monkeys in our *in situ* attachment model. The virulence between DS17 and its papG mutant DS17-8 in the cynomolgus monkey was compared. To study the role of the PapG adhesin in pyelonephritis, five monkeys were infected with *E. coli* strain DS17 whereas six monkeys received mutant strain DS17-8 via a cytoscopically inserted ureteral catheter, and it was shown that there was a significant difference between the two groups. The monkeys receiving wild type strain DS17 has a mean bacteruria of 21 days compared to 6.8 days for those monkeys receiving the mutant strain DS17-8. Renal clearance was also significantly lower for the wild type strain and renal function was significantly reduced in the monkeys receiving the wild type strain but not in those monkeys receiving the isogenic papG mutant strain. Pathologic evaluation of the infected kidney confirmed the functional studies showing significantly less inflammation and pathologic changes in the mutant group as compared to those receiving the wild type.

Therefore, it is concluded that the Gal $\alpha$ 1-4Gal binding PapG adhesin at the tip of P pili is required for pyelonephritis to occur in the normal urinary tract of primates. Up to now, no such direct demonstration for the role of a pilus associated adhesin in a specific bacterial infection has been made.

Interestingly, in the experiments described above, no evidence was found that the PapG adhesin was required for colonization of the lower urinary tract or for the development of acute cystitis. Both strains colonized the vagina, persisted

in the intestine and caused bladder infection. Thus, although *PapG* was a critical virulence determinant in causing pyelonephritis, it did not appear to be crucial in cystitis. However, this did not exclude the possibility that other  
5 components of the P pilus or that another type of pilus, such as type 1, was required for cystitis.

These hypotheses will be investigated by testing the virulence of isogenic DS17 mutant strains containing site directed mutations in Arg-8 of *PapD* or *FimC*. The mutants will  
10 presumably be defective in their ability to assemble P and type 1 pili. If the *papD* mutant is non-virulent in the monkey cystitis model, it may indicate that some component of the P-pilus other than *PapG* is essential for *E. coli* to generate bladder infection. For example, the major component of the  
15 tip fibrillum, *PapE*, has been shown to bind fibronectin and fibronectin could be an important factor in cystitis. Similarly, if the *fimC* mutant is non-virulent, it would confirm the role of the mannose-binding type 1 pili in causing cystitis. The ability of these point mutations to abolish the  
20 virulence of DS17 in causing either pyelonephritis or cystitis would validate the therapeutic potentials for a chaperone inhibitor.

#### EXAMPLE 9

25 Identification of the motif of binding between *PapD* and K1'-19'WT

In order to confirm the generality of the binding mode for C-terminal peptides of pilin subunits to *PapD* as observed in the *PapD*-G1'-19'WT crystal structure (see Example 1) a second *PapD*-peptide complex was investigated by X-ray crystallo-  
30 graphy. Since tight binding had been observed for the peptide derived from the wild-type C-terminal 19 amino-acids of *PapK* (K1'-19'WT, SEQ ID NO: 18, and numbered here from the C-terminal Arg-1' to the N-terminal Lys-19') it was decided to use this as the peptide for the second *PapD* complex.

*Obtaining material; protein & peptide.*

- PapD was prepared as previously described (Holmgren et al., 1988) and obtained from Dr. Scott Hultgren, Dept of Molecular Microbiology, Washington University School of Medicine, St. Louis USA. The peptide K1'-19'WT was prepared by Fmoc solid phase synthesis, purified by reversed phase HPLC and obtained from Dr. Jan Kihlberg, Dept. of Chemistry, University of Lund, Lund, Sweden.

*Crystallisation of PapD-peptide complex.*

- 10 After a number of different experimental conditions had been explored around those previously used to obtain PapD-G1'-19'WT crystals the best crystals of the PapD-K1'-19'WT complex were grown by vapour diffusion against 20% PEG8000, 0.1 M MES pH 6.5. The crystallisation drop  
15 contained equal volumes of reservoir and protein solution. The protein solution (15 mg/ml) contained a 1:1 molar ratio of PapD to peptide in 20 mM MES pH 6.5 with 1.0%  $\beta$ -octyl glucoside ( $\beta$ -OG).

- These crystals were mounted inside sealed quartz-glass capillary tubes and initially characterised by examining them on a  
20 X-ray precession camera. From standard analysis of such images it was determined that the above-mentioned crystals have a orthorhombic space group, C2221 (thus differing from the PapD-G1'-19'WT crystals which were C2), with cell dimensions  
25  $a=57.1 \text{ \AA}$ ,  $b=153.2 \text{ \AA}$ ,  $c=135.4 \text{ \AA}$  and  $\alpha=\beta=\gamma=90^\circ$ , 2 molecules in the asymmetric unit and diffract to  $2.7 \text{ \AA}$  resolution on a lab X-ray source with rotating anode and Cu  $K_\alpha$  target.

*Collection & processing of experimental data.*

- The intensity data for the PapD-K-peptide crystals were  
30 collected on a R-AXIS II area-detector system (from R-AXIS) at Symbicom AB., Uppsala. All data were obtained from a single crystal and processed initially with the DENZO soft-



ware package (Otwinosky, 1993). Merging and scaling of the data, however, was carried out using ROTAVATA and AGROVATA from the CCP4 package (CCP4, 1979). The final data set contained 15,989 independent reflections with an Rsym of 8.7% for data between 20.0 and 2.7 Å resolution.

#### *Solution of three dimensional structure.*

The structure of the complex was solved by the standard method of molecular replacement using the program XPLOR (Brunger, 1992). The search model used was the refined 2.0 Å resolution structure of *PapD* (Holmgren and Bränden, 1989). Using 8.0 to 4.0 Å resolution data the self-rotation function again gave a clear non-crystallographic two-fold axis. The top peaks in the translation functions also gave the correct solutions. After the translation functions the R-factor was 36.7% for 8.0 to 4.0 Å resolution data. Subsequent rigid body refinement in which all 4 domains of the 2 *PapD* molecules in the asymmetric unit were allowed to refine independently resulted in an R-factor of 33.6% for the same data.

Examination of an  $|F_o| - |F_c|$  electron density map at this stage using the graphics program O (Jones and Kjeldgaard, 1994) showed clear density corresponding to the K-peptide in the *PapD* cleft and running along the surface of the protein in an analogous fashion to that found for the G-peptide. The orientation of the peptide was easily determined from the electron density, but initially only the final 12 C-terminal amino-acids of the peptide could be modelled into density.

#### *Refinement and analysis of structure.*

Simulated annealing refinement with XPLOR (Brunger, 1992) was initiated at this stage. Several additional cycles of model building and refinement were carried out with a further 2 peptide amino-acids being added to the N terminal end of the peptide to yield an R-factor for the current model of 19.2% for 8.0 to 2.7 Å resolution data. The model at the present

stage of refinement (which contains no water molecules and does not include the first 5 N-terminal amino-acids of the peptide) has root-mean square (rms) deviations from ideal geometry of 0.019 for bonds lengths and 3.8° for bond angles.

- 5 Despite the two *PapD*-peptide complexes being solved in different space groups, their overall structures are found to be highly similar with the peptide interacting with *PapD* in a manner essentially identical to that previously seen for the *PapD*-G1'-19'WT structure. Thus the K1'-19'WT peptide is again  
10 seen to bind in an extended conformation with the C-terminal Arg-1' anchored within the inter-domain cleft and subunit binding site. Hydrogen bonds are also formed between the peptide carboxy terminus and two invariant positively charged residues of *PapD*, Arg-8 and Lys-112. The K1'-19'WT peptide  
15 then run along the surface of the N-terminal domain, forming a parallel  $\beta$ -strand interaction with strand G1. In this way between 9 mainchain hydrogen bonds are formed between residues 10' to 2' of the peptide and Val-102 to Lys-110 of *PapD* and thus extending the  $\beta$ -sheet of *PapD* out into the peptide.
- 20 In addition, a dimer association is observed between the two *PapD*-peptide complexes within the unit cell similar to that seen in the *PapD*-G1'-19'WT crystal structure. Again the  $\beta$ -sheet is extended as a result of a non-crystallographic 2-fold symmetry which places a second *PapD*-peptide complex  
25 adjacent to the first such that the two bound peptide chains interact as anti-parallel  $\beta$ -strands with a mixed  $\beta$ -sheet again being created between the two complexes involving a total of 10  $\beta$ -strands. The only major difference between the two *PapD*-peptide complexes is the fact that in the  
30 *PapD*-K1'-19'WT structure the peptides of the non-crystallographically related complexes are positioned two residues closer to the COOH-terminus of its partner. Thus while eight hydrogen bonds are formed between the peptides in the *PapD*-G1'-19'WT complex a total of 10 are observed in the  
35 *PapD*-K1'-19'WT.

Apart from the C-terminal residues Tyr-2' and Arg-1' there are again relatively few contacts between the sidechains of the peptide and PapD. The major interactions are provided by the mainchain hydrogen bonds to the G1 strand. There are, however, a number of hydrophobic interactions within the  $\beta$ -sheet, in particular between the peptide's Tyr-6' with Ile-105 and Leu-107 of strand G1.

#### EXAMPLE 10

Identification of a second binding site in PapD and development of new assays

##### Introduction

A chaperone binding assay was developed to delineate PapD-PapG interactions using MBP/G fusions, containing increasing lengths of the COOH-terminus of PapG. The ability of PapD to bind to PapG truncate proteins missing increasing lengths of the COOH terminus of PapG was also examined.

##### Experimental procedures

##### Bacterial strains

The *E. coli* strain HB101 (Maniatis et al., 1982) was used as a host strain in the studies involving the MBP/G fusion proteins. Strain DH5 $\alpha$  was used as the host to construct MBP/G fusion proteins (Hanahan, 1983) and PapG truncates. KS474 (*degP::kan*) (Strauch et al., 1989) was kindly provided by J. Beckwith and used for expression of the PapG truncate proteins.

##### Plasmid construction

The *E. coli* expression plasmid pMAL-p2 (new England Biolabs, Beverly, MA, USA) was used for the construction of pMAL-p4, pMAL/G1'-19', pMAL/G1'-81' and pMAL/G1'-140' using procedures

essentially from Maniatis (Maniatis et al. 1982). The *malE* gene encoded by pMAL-p2 has a *lacZ $\alpha$*  sequence fused to the 3' end. Plasmid pMAL-p2 was digested with *SalI* and filled in with Klenow (Maniatis et al., 1982). The fragment was re-

5 ligated and the resulting plasmid called pMAL-p4 had a stop codon between the *malE* and *lacZ $\alpha$*  sequences. Plasmid pMAL-p4 was digested with *HindIII* and filled in with Klenow and the resulting fragment was ligated with a *ClaI* linker, 5'CCATCGATGG3' (New England Biolabs, Beverly, MA, USA) to

10 produce plasmid pMAL-p5. Primer 23969 (5'CCCCCTGCAGATCAGATTAAGCAGCTACCTGCG3', SEQ ID NO: 23) and primer 23557 (5'CCCCTGCAGTAAAAATATCTCTGCTCAGAAATAC3', SEQ ID NO: 24) and primer 23559 (5'CCATCGATGAACAGCCAGTCAGATAATC3', SEQ ID NO: 25) were used for Polymerase Chain Reactions (PCR)

15 utilizing pPAP5 (Hull et al., 1981; Lindberg et al., 1984) as the template. The DNA fragments encoding the region of amino acid residues 1'-81' and 1'-140' of the *PapG* were obtained from the PCR reactions using primers 23557 and 23559; and primers 23969 and 23559, respectively. The amplified DNA

20 fragments were purified and restricted with *PstI* and *ClaI* and ligated into *PstI* and *ClaI* restricted pMAL-p5 fragment to construct pMAL/G1'-81' and pMAL/G1'-140', respectively. To create pMAL/G1'-19', two nucleotide oligomers (5'GGAAAGAGAAAACCCGGGGAGCTATCTGGTTCTATGACTATGGTTCTGAGTTTCCCCT

25 GAT3', SEQ ID NO: 26 and 5'CGATCAGGGGAAACTCAGAACCATAGTCATAGAACCAGATAGCCCCGGGTTTTCTCTTT CCTGCA3', SEQ ID NO: 27) were chemically synthesized and annealed. The annealed DNA fragment containing the sequence of the carboxyl terminal 1'-19' amino acid residues of *PapG*

30 was ligated to *PstI*-*ClaI* pMAL-p5 fragment to produce pMAL/G1'-19'. All the constructs used in this study were confirmed by DNA sequencing analysis using Sequenase Version 2.0 according to the manufacturer's directions (USB, Cleveland, OH, USA). pHJ14 (*PapG*3) was created by digestion of

35 pHJ8 (wt *PapG* in *P<sub>tac</sub>* promoter plasmid, pMMB66) with *BamHI* and *BglIII*, removing approximately 500 nucleotides, purification of the large fragment and religation. pHJ23 (*PapG* 2) was created by first cloning *papG* into pUC19 at *EcoRI* and *BamHI*,

digesting with *HincII* to remove the last approximately 300 nucleotides and religating. The *EcoRI-HincII* fragment was then cloned into pMMB66 to create pHJ23.

#### *Induction and partial purification of MBP/G fusions*

5 Strains carrying MBP or MBP/G fusion gene plasmids were induced in LB broth culture with 1 mM isopropylthiogalactoside (IPTG). Periplasmic extracts were prepared as described (Slonim et al., 1992). One tenth volume of 10X phosphate-buffered saline (PBS, 120 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>,  
10 and 2 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4) was added to the periplasmic extract. Amylose resin was then added to the periplasmic extracts at a 1:5 ratio. The mixture was rocked at 4°C overnight and the beads were subsequently washed 5 times with PBS. MBP or MBP/G fusion proteins were eluted with 20 mM  
15 maltose in PBS by rocking at 4°C for an hour. Protein concentrations were determined using the Bio-Rad DC protein assay kit (Bio-Rad Laboratories, Hercules, CA, USA). The full length fusion protein concentration was further quantitated by the Coomassie stained SDS-PAGE gel with known concentrations  
20 of bovine albumin (BSA) as the standards followed by densitometry scanning.

#### *Characterization of interactions between PapD and MBP/G fusion proteins*

The MBP or MBP/G fusion preparations from HB101/pLS101/pMAL-p4, HB101/pLS101/pMAL/G1'-19', HB101/pLS101/pMAL/G1'-81' and  
25 BH101/pLS101/pMAL/G1'-140' were applied to isoelectric focusing (IEF) pI 3-9 gels (Pharmacia Phast System, Pharmacia, Sweden) or 12.5% SDS-PAGE followed by silver staining, Coomassie blue staining or immunoblotting with anti-PapD  
30 antiserum as described (Kuehn, et al., 1991).

The stability of the PapD-MBP/G1'-140' complex in urea was determined by incubating 1 µg partially purified MBP/G1'-140' preparation from HB101/pLS101/pMAL-G140 in 0.5 M urea for 5

minutes at 25°C (Kuehn et al., 1991). The proteins were analyzed by IEF (pI 3-9) with silver staining. The PapD-MBP/G1'-140' complex on the IEF gel was quantified by densitometry (Digital Imaging System, Is-1000).

- 5 The interaction between PapD and MBP/G fusion *in vitro* was assayed by incubating purified PapD (1 µg) and amylose bead affinity purified MBP/G fusion protein (1 µg) in 4 µl PBS for 30 minutes and then applied to IEF (pI 3-9), followed by silver staining. Another *in vitro* assay that is more quantitative is the ELISA described below.
- 10

#### *Characterization of the interaction between PapD and PapG truncates*

- PapD and the PapG-truncates were co-expressed in KS474, the truncates were induced at OD<sub>600</sub> 0.6 with 0.5 mM IPTG and PapD under the control of the arabinose promoter was induced with 0.2% arabinose. After one hour induction, periplasmic extracts were prepared and rocked with Gal α(1-4) Gal beads overnight at 4°C and eluted with 15 µg/ml Gal α(1-4) Gal-TMSET solution as described previously (Hultgren et al., 1989; Kuehn et al., 1991). The eluates were analyzed by acidic native gel electrophoresis (Striker et al., 1994) followed by western blotting using anti-PapD and anti-PapG antiserum.
- 15
- 20

#### *Hemagglutination titers*

- 25 HB101/pFJ3 harbouring various MBP/G fusion plasmids was passaged on Trypticase Soy Agar (TSA, Becton Dickinson, Cockeysville, MD, USA) three times to induce P-pili expression. At the last passage, the expression of MBP/G fusions was induced on TSA containing 100 µM IPTG. The cells from 30 different strains were then collected and the HA titers were determined as described (Jacob-Dubuisson et al., 1993b).

*Pili preparation and quantitation*

The cells obtained from TSA plates were collected. Pili were prepared from the same amounts of cells using the procedure described previously (Jacob-Dubuisson, et al., 1993b). These 5 pili preparations were boiled in Lammeli sample buffer with 4 M urea, then analyzed on Coomassie blue stained SDS polyacrylamide gels. The relative amount of piliation was quantified by densitometric scanning of the PapA bands.

*ELISA assay*

- 10 Stock solutions of MBP/G fusion proteins in PBS were diluted to 40 pmol/50  $\mu$ l with PBS. The solutions of the 40 pmol/50  $\mu$ l MBP/G fusion proteins were serially diluted in PBS and 50  $\mu$ l of each solution were used to coat the wells of a 96-well Nunc Immunoplate (Inter Med, DK-4000 Roskilde, Denmark)
- 15 overnight at 4°C. The wells were then washed with PBS and blocked with 200  $\mu$ l of 3% BSA in PBS for 2 hours at 25°C. The plates were washed vigorously three times with PBS and incubated with 50  $\mu$ l of PapD diluted at 50 pmol PapD in 50  $\mu$ l 3% BSA-PBS for 45 minutes at 25°C. After three washings with
- 20 PBS, the wells were incubated with a 1:500 dilution of rabbit anti-PapD antiserum in 3% BSA-PBS for 45 minutes at 25°C. After three washings with PBS, the wells were incubated with a 1:1000 dilution of goat antiserum to rabbit IgG coupled to alkaline phosphatase in 3% BSA-PBS for 45 minutes at 25°C.
- 25 After three washings with PBS and three washings with developing buffer (10 mM diethanolamine, 0.5 mM MgCl<sub>2</sub>), 50  $\mu$ l of filtered 1 mg/ml p-nitrophenyl phosphate (Sigma, St. Louis, MO, USA) in developing buffer was added. The reaction was incubated for 1 hour in the dark at 25°C and the absorbance
- 30 at 405 nm was read.

For assaying the second site peptides, all the peptides were dissolved in dimethyl sulfoxide (DMSO, Sigma, St. Louis, MO, USA) to a final concentration 2 mM. The stock solutions were diluted to a proper concentration in PBS and coated overnight

on microtiter wells at 4°C. All the peptide solutions were adjusted to the same DMSO concentration. The subsequent steps followed the procedure described above.

## Results

### 5 MBP/G fusions

To identify determinants on *PapG* essential for *PapD* interaction we constructed three in frame fusions between the *MalE* gene, encoding maltose binding protein, and sequences encoding the COOH-terminal residues 1'-19', 1'-81' and 1'-140' of  
10 *PapG* (Fig. 20). The resulting chimeric proteins were called MBP/G1'-19', MBP/G1'-81' and MBP/G1'-140', respectively, to indicate the regions of *PapG* present in the MBP/G fusions. MBP/G1'-19' was created since *PapD* was recently shown to bind a peptide consisting of the C-terminal 19 amino acids of *PapD*  
15 in vitro (Kuehn et al., 1993). MBP/G1'-81' and MBP/G1'-140' were created to test the requirement of the disulfide bridge (C196-C228 in *PapG*) for recognition by *PapD*. A similarly located disulfide bond is found in virtually all pilus subunits (Simons et al., 1990). MBP/G1'-81' lacks the two  
20 cysteine residues while MBP/G1'-140' contains the disulfide bond (Fig. 20).

### *PapD*-MBP/G interactions in vivo

The ability of *PapD* to bind to the MBP/G fusions was investigated using amylose affinity chromatography (Kellermann and  
25 Ferenci, 1982). *PapD* was co-expressed from plasmid pLS101 (Slonim et al., 1992) with each of the MBP/G fusions. Periplasmic extracts from each strain containing the MBP/G proteins and *PapD* were subjected to amylose affinity chromatography and the eluates analyzed by SDS-PAGE (Fig. 21A) and by  
30 western blotting with anti-*PapD* antisera (Fig. 21B). In this assay, co-elution of *PapD* signified the ability of *PapD* to interact with the MBP/G protein. The western blot revealed that *PapD* co-eluted with all three fusion proteins (Fig. 21B,



lane 2, 3 and 4) but not with the MBP control (Fig. 21A and 21B, lane 1). However, *PapD* interacted much stronger with the MBP/G1'-140' fusion as can be seen in the western blot (Fig. 21B, lane 4) as well as in the coomassie blue stained gel of the eluates (Fig. 21A, lane 4). Thus, *PapD* interacted strongly with the MBP/G1'-140' protein and only weakly with the MBP/G1'-19' and MBP/G1'-81' proteins.

Analysis of the eluates on silver stained isoelectric focusing (IEF) gels revealed that the *PapD*-MBP/G1'-140' complex migrated at an isoelectric point (pI) of 5.2 (Fig. 21C, lane 4) which was intermediate between the pIs of MBP/G1'-140' (~4.4) and *PapD* (~9.1). This band was confirmed to contain both *PapD* and MBP/G1'-140' by excising the unstained band, applying the material to SDS-PAGE and analyzing it by western-blotting using anti-*PapD* and anti-MBP antisera (Fig. 21D). It was not possible to detect stable complexes in the eluates of MBP, MBP/G1'-19' or MBP/G1'-81' when co-expressed with *PapD* (Fig. 21C, lane 1, 2 and 3).

The stability of the *PapD*-MBP/G1'-140' complex was measured as a function of urea concentration in the presence of 15 mM DTT. Both the *PapD*-*PapG* and *PapD*-MBP/G1'-140' complexes behaved similarly under these conditions and were dissociated after incubation in 2M urea (data not shown). By these criteria, the *PapD*-MBP/G1'-140' complex was as stable as the *PapD*-*PapG* complex.

#### *Expression of MBP/G proteins inhibit pilus formation*

*PapD* is essential for P-pilus assembly (Hultgren et al., 1991). A decrease in the concentration of *PapD* in the periplasm has been shown to cause a concomitant decrease in piliation (Slonim et al., 1992). The ability of the MBP/G fusions to block pilus formation by inhibiting chaperone-subunit complex formation when co-expressed in trans with the *pap* operon was tested. Plasmids pMAL-p4, pMAL/G1'-19', pMAL/G1'-81' and pMAL/G1'-140' were transformed into the

strain HB101/pFJ3 which contains the *pap* operon under the control of its own promoter in the vector pACYC184 (Jacob-Dubuisson et al. 1993b). Each MBP/G fusion protein was localized in the periplasm as determined by SDS-PAGE and western blotting with anti-MBP antisera (data not shown). Hemagglutination (HA) assays were performed on the cells co-expressing the *pap* operon with either MBP (pMAL-p4) or with each of the MBP/G fusions. Co-expression of the MBP/G1'-140' fusions with the *pap* operon decreased the HA titer 30 fold compared to the MBP control (see table 3).

TABLE 3: MBP/G fusions inhibit P pilus formation

MBP/ proteins <sup>a</sup>	MBP	MBP/G1'-19'	MBP/G1'-81'	MBP/G1'-140'
HA titer <sup>b</sup>	128	64	82	4
Piliation percentage <sup>c</sup>	100%	50%	40%	10%

<sup>a</sup> Plasmids pMAL-p4, pMAL/G1'-19', pMAL/G1'-81' and pMAL/G1'-140' encoding MBP, MBP-G1'-19', MBP-G1'-81' and MBP-G1'-140', respectively, were transformed into the strain HB1010/pFJ3 and the expression of *pap* operon and each of the MBP/G fusions was induced in each strain.

<sup>b</sup> Highest dilution of cell suspension yielding detectable HA

<sup>c</sup> The pili preparations were analyzed on Coomassie blue stained gels and the *PapA* bands were scanned by a densitometer. The relative piliation was determined by equating the densitometric value of *PapA* to 100% of the control. (HB101/pFJ3/pMAL-p4).

Co-expression of the MBP/G1'-81' and MBP/G1'-19' fusions with the *pap* operon had only weak effects on the HA titer (Table 3). Co-expression of MBP/G1'-140' with the *pap* operon reduced the amount of pili that could be purified from the cells by 90% whereas MBP/G1'-19' and MBP/G1'-81' reduced the amount of pili formed by about 50% compared to MBP alone (Table 3). Electron microscopy confirmed that cells expressing the MBP/G1'-140' fusion had little or no pili compared to the fully piliated cells co-expressing MBP (data not shown).

We hypothesized that the co-expression of the MBP/G1'-140' fusion with the *pap* operon inhibited pilus formation by

titrating *PapD* away from the subunits thus driving subunits down dead end pathways of aggregation and proteolytic degradation (Hultgren et al., 1989; Holmgren et al., 1992). This hypothesis was tested as follows. pMAL-p4, pMAL/G1'-19',  
5 pMAL/G1'-81' or pMAL/G1'-140' were transformed into BH101 carrying pFJ22 (Jacob-Dubuisson et al., 1994). Plasmid pFJ22 encodes *papDJKEFGA* under the control of the  $P_{tac}$  promoter. In HB101/pFJ22, pilus subunits accumulate in the periplasmic space due to the absence of the usher, *PapC*. The effect of  
10 expressing each fusion on the fate of each pilus subunit type was determined by western blotting (Fig. 22). A significant decrease in the amount of *PapG*, *PapA* and *PapF* in cells co-expressing the MBP/G1'-140' fusion was detected arguing that the MBP/G1'-140' fusion blocked chaperone-subunit complex  
15 formation by interacting with *PapD* and trapping it away from pilus subunits. There was no effect on *PapK* and *PapE* stability for reasons that are not understood. Nevertheless, the fusion was capable of blocking pilus formation by interfering with the formation of critical chaperone-subunit complexes.

## 20 *In vitro* chaperone binding assays

*PapD* binding to the MBP/G fusion proteins was further investigated using two different *in vitro* assays. In the first assay, purified MBP/G fusion proteins or MBP, were incubated with *PapD* and complex formation was analyzed on silver  
25 stained IEF gels. *PapD* bound to the MBP/G1'-140' fusion protein and formed a stable complex that migrated to an identical pI (5.2) as that seen for the complex formed *in vivo* (Fig. 23A, lane 4). In contrast, no complexes between *PapD* and MBP/G1'-81', MBP/G1'-19' or the MBP control were  
30 detected (Fig. 23A, lane 1, 2 and 3). In the second assay, we investigated the ability of *PapD* to bind to the three different MBP/G fusion proteins immobilized on microtiter wells and quantitated the interactions in an enzyme-linked immunosorbent assay (ELISA) using anti-*PapD* antiserum (Fig. 23B). *PapD*  
35 bound weakly to MBP/G1'-19', slightly better to MBP/G1'-81' but very strongly to the MBP/G1'-140'. *PapD* did not bind the

MBP control. The dramatic increase in the affinity of PapD for the MBP/G1'-140' fusion as compared to the MBP/G1'-81' and MBP/G1'-19' fusions, suggested that residues 81' to 140' in PapG were critical for strong PapD binding.

- 5 The invariant cleft residues Arg-8 and Lys-112 in PapD have been shown to form a molecular anchor in the chaperone cleft necessary for binding pilus subunits (Slonim et al., 1992; Kuehn et al., 1993). Mutations in these residues abolished or greatly reduced the ability of PapD to bind subunits and  
10 mediate pilus assembly (Kuehn et al., 1993). Arg-8A and Lys-112A mutant PapD proteins demonstrated a greatly reduced ability to bind to the MBP/G1'-140' fusion protein both in vivo and in vitro (data not shown). These results argue that the interactions between PapD and the MBP/G1'-140' fusion are  
15 biologically relevant.

*The upstream site is an independent binding determinant*

- We tested the ability of PapD to bind three PapG COOH-terminal truncate proteins (shown in Fig. 20) in order to delineate the limits of the upstream PapD binding site and to  
20 test whether it was capable of functioning as an independent site. The truncate proteins were co-expressed with PapD in strain KS474 (Strauch et al., 1989) which carries a kanamycin cassette in the *degP* locus (*degP41* mutant). Both the PapG2 and PapG3 truncates were expressed and stable in KS474,  
25 however the PapG1 truncate (removing the last 14 residues) underwent limited degradation (data not shown). Complex formation was assayed by analyzing periplasmic extracts on acid native polyacrylamide gels (Striker et al., 1994). Anti-PapD and anti-PapG antisera was used to detect the PapD-PapG  
30 complexes after western blotting (Fig. 24). PapD-PapG complexes are resolved from PapD alone on these gels due to their differences in size, shape and charge (Striker et al., 1994) PapD formed a complex with full-length PapG as well as with the PapG2 truncate (Fig. 24A, lanes 2 and 3). The com-  
35 plex band was also recognized by anti PapG antiserum (Fig.

24B, lanes 2 and 3). *PapD* did not form a complex with the *PapG3* truncate (Fig. 24A & B, lane 4) which terminates at amino acid 145. This information delineates an endpoint for the second *PapD* binding site (as indicated in Fig. 20).

- 5 Combining the *PapG* truncate and the MBP/G fusion data (shown in Fig. 20) suggested that the second site on *PapG* recognized by *PapD* resided in a region between residues 117' to 141' of *PapG*. Four overlapping peptides were synthesized corresponding to the region between residues 120' to 156' and tested  
10 for their ability to bind *PapD* in an ELISA assay. Such a strategy proved successful in studying the COOH-terminal *PapD* binding site (Kuehn et al., 1993). The peptide corresponding to residues 125' to 140', but not the other three peptides, bound to *PapD* in the ELISA (Fig. 25). These data argue that  
15 *PapD* recognizes two surfaces on *PapG*. *PapD* forms a beta strand zippering interaction with the COOH-terminus but also recognizes a region containing residues 125'-140'.

#### *Discussion and conclusions*

- In conclusion, a chaperone binding assay was developed using  
20 fusions of the carboxyl terminus of *PapG* to Maltose Binding Protein (MBP/G fusions) to investigate whether chaperone-subunit complex formation requires additional interactions. *PapD* bound strongly to an MBP/G fusion containing the C-terminal 140 amino acids of *PapG* (MBP-G1'-140') but only weakly  
25 to the MBP-G1'-81', arguing that the region between the C-terminal residues 81' and 140' contains additional information that is required for strong *PapD*-*PapG* interactions. *PapD* was further shown to interact with a *PapG* C-terminal truncate containing residues 117'-314' (corresponding to a  
30 truncate consisting of the first 198 N-terminal amino acid residues of *PapG*), but not with a truncate containing residues 170'-314' (corresponding to a truncate consisting of the first 145 N-terminal amino acid residues of *PapG*).

- These results together suggest that a second, independent *PapD* interactive site exists which does not interact with the C-terminus of the pilus subunit. This assumption is further confirmed by the fact that one of four peptides overlapping the second site region of *PapG* was recognized by *PapD*. The last result also rules out that the function of the upstream site was to exert secondary effects on the COOH terminal chaperone binding site since this region was independently capable of facilitating *PapD* binding.
- 10 Unlike what is true for the native subunits it was demonstrated that MBP-G1'-140' is capable of folding even in the absence of *PapD* and remain soluble in the periplasmic space; this folding involves the formation of an intramolecular disulphide bond. This is consistent with earlier data showing
- 15 that the subunits themselves are highly folded when bound to *PapD* (Kuehn et al., 1991; Striker et al., 1994). However, it is not yet known whether *PapD* binds to subunits *in vivo* after they fold or whether folding occurs in the context of an interaction with *PapD*.
- 20 Finally, the expression of the MBP/G fusions in cells producing P pili inhibited pilus assembly to varying degrees by binding to *PapD* and preventing chaperone-subunit complex formation. The ability of the fusions to block pilus assembly increased as the length of the fusion protein increased.
- 25 MBP-G1'-140' was a strong inhibitor of pilus assembly while MBP-G1'-81' and MBP-G1'-19' were weak inhibitors. This is consistent with the finding that residues 140'-81' contain a region necessary for strong interactions between *PapD* and *PapG*.
- 30 Recent research by the inventors has further revealed that the simultaneous expression *in vivo* of *PapG* and a modified *PapD* polypeptide free of domain 1 in a *degP41* strain (KS474) expressing this truncated *PapD* results in 1) suppression of *PapG* toxicity in this cell line and 2) efficient partitioning

into the periplasmic space of PapG. However, the pilus subunit is not correctly folded after release from PapD.

It is thus suggested by the inventors, that the binding between pilus subunits and PapD takes place by the pilus subunit binding to domain 1 and 2 of PapD. The binding between the subunit and domain 1 is important for the correct folding of the subunit, and the binding between the subunit and domain 2 is important for the transport of the subunit out into the periplasm. Whether domain 2 of PapD also participates in the folding of subunits is not known.

Since the above results are strong indicators of the existence of at least one binding site apart from the one involving Arg-8 and Lys-112 of PapD, and since it seems that this novel binding site is also very important in the net interaction between PapD and the pilus subunits, it is contemplated that also effecting this binding site will have the antibacterial effects described herein.

It is thus the plan to elucidate the motif of binding between this second binding site of PapD in manners similar to those described herein, and it is further the plan to design/identify compounds capable of interacting with this second binding site in order to ultimately synthesize compounds capable of interacting with this site in such a manner that assembly of intact pili is prevented, inhibited or enhanced.

## REFERENCES

- Allen B L et al., 1991, *J. Bacteriol.*, 173, 916-920.
- Amit A G et al., 1988, *Science*, 230, 747-753.
- Baga M et al, 1987, *Cell*, 49, 241.
- 5 Bakker et al., 1991., *Molec. Microbiol.*, 5, 875.
- Ben-Naim and Marcus, 1984, *J. Chem. Phys.*, 81, 2016-2027.
- Bertin Y et al., 1993, *FEMS Microbiol. Lett.*, 108, 59.
- Blond-Elguindi S et al., 1993, *Cell*, 75, 717-728.
- Bock K et al., 1985, *J. Biol. Chem.*, 260, 8545-8551.
- 10 Boobbyer, 1989, *J. Med. Chem.*, 32, 1083-1094.
- Bothner-By A A, 1965, *Adv. Magn. Resonance*, 1, 195
- Brint A T and Willet P, 1987, *J. Mol. Graphics*, 5, 49-56.
- Brunger A T, 1992, *X-PLOR Manual*, Version 3.0, Yale Univ., New Haven.
- 15 Burkert U and Allinger N L, 1982, *Molecular Mechanics*, ACS Monographs, Washington D.C.
- CCP4, The SERC, UK Collaborative Project No. 4, A Suite of Programs for Protein Crystallography, Daresbury Laboratory, UK.
- 20 Chandler D et al., 1983, *Science*, 220, 787-794.
- Clouthier S C et al., 1993, *J. Bacteriol.*, 175, 2523.



- Dodson et al., 1993, *Proc. Natl. Acad. Sci.*, 90, 3670-3674.
- Galyov E E et al., 1991, *FEBS Lett.*, 286, 79.
- Gerlach G D, Clegg S. and Allen B L, 1989, *J. Bacteriol.*, 171, 1262-70.
- 5 Gething M-J and Sambrook J, 1992, *Nature*, 355, 33.
- Goodford, 1985, *J. Med. Chem.*, 1985, 28, 849-857.
- Grant G A et al., 1992, *Synthetic Peptides, A Users Guide*, W. A. Freeman and Company, New York.
- 10 Hanahan D, *J. Mol. Biol.*, 1983, 166, 557-580.
- Hirschmann R et al, 1992, *J. Am. Chem. Soc.*, 114, 9217-9218.
- Holmgren A et al., 1988, *J. Mol. Biol.* 203, 279.
- Holmgren A and Bränden C I, 1989, *Nature*, 342, 248.
- Holmgren A et al., 1992, *The EMBO Journal*, 11, 4, 1617-1622.
- 15 Hultgren S J et al., 1989, *Proc. Natl. Acad. Sci. USA*, 86, 4357.
- Hull R A et al., 1981, *Infect. and Immun.*, 33, 933-938.
- Hultgren S J et al., 1993, *Cell*, 73, 887-901.
- Hultgren S J et al., 1989, *Proc. Natl. Acad. Sci. USA*, 86,  
20 4357-61.
- Hultgren S J et al., 1991, *Annu. Rev. Microbiol.*, 45, 383-415.
- Hultgren S J et al., 1993, *Cell*, 73, 887-901.

- Hällgren C and Widmalm G, 1993, *J. Carbohydr. Chem.*, 12(3), 309-333.
- Iriarte et al., 1993, *Mol. Microbiol.*, 9, 507.
- Jacob-Dubuisson F et al., 1993a, *Trends Microbiol.*, 1, 50-55.
- 5 Jacob-Dubuisson F et al., 1993b, *EMBO J.*, 21, 837-847.
- Jacob-Dubuisson F et al., 1994, *J. Biol. Chem.*, 269, 12447-12455.
- Jalajakumari M B et al., 1989, *Mol. Microbiol.*, 3, 1685.
- Jones A T and Kjeldgaard M, 1994, O Version 5.9, Dept. Molec.  
10 Biol. Uppsala. Sweden.
- Jorgensen W L, 1986, *J. Phys. Chem.*, 90, 1276-1284.
- Kawaminami M et al, 1981, *Acta Crystallogr. Section B* 107(37), 2026.
- Kellermann O K and Ferenci T, 1982, *Methods in Enzymology*,  
15 90, 459-463.
- Klann A G et al., 1994, *J. Bacteriol.*, 176, 2312-2317.
- Klemm P et al., 1992, *Res. Microbiol.*, 143, 831.
- Kuehn M. J., Normark S, and Hultgren S J., 1991. *Proc. Natl. Acad. Sci.* 88, 10586-10590.
- 20 Kuehn M.J. et al, 1993, *Science*, 262, 1234-1241.
- Lam K T and Calderwood S K, 1992, *Biochem and Biophys., Res. Comm.* 184, 167.
- Landry S J and Gierasch L M, 1991, *Biochemistry*, 30, 7359.

- Landry S J et al., 1992, *Nature*, 355, 455.
- Lee F S et al., 1992, *Prot. Eng.*, 5, 215-228.
- Lindberg F et al., 1984, *EMBO J.*, 3, 1167-1173.
- Lindberg F et al., 1987, *Nature*, 328, 84-87.
- 5 Lindberg F et al, 1989, *J. Bacteriol.* 171, 6052.
- Lindler et al., 1993, Genbank, accession no: M86713.
- Lintermans P, 1990, Thesis, Rijksuniversiteit Ghent, Belgium.
- Locht C et al., 1992, *EMBO J.*, 11, 3175.
- Lund B et al., 1987, *Proc. Nat. Acad. Sci. USA*, 84, 5898.
- 10 Marcus R A, 1964, *Ann. Rev. Phys. Chem.*, 15, 155-196.
- Maniatis T et al., 1982, *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, NY.
- Messerschmit A and Pflugrath J W, 1987, *J. Appl. Crystallogr.*, 20, 306.
- 15 Normark S et al., Genetics and Biogenesis of *Escherichia coli* Adhesin. In *Microbial lectins and Agglutinins: Properties and Biological Activity*, D. Mirelman (ed.), Wiley Interscience, New York, 113-143.
- Otwinowski Z, 1993, *Proc. CCP4 Study Weekend 29-30 Jan 1991*.
- 20 Data Collection and Processing. SERC Daresbury Lab.
- Patroni J J et al, 1988, *Aust. J. Chem.*, 41, 91-102.
- Pratt L R and Chandler D, 1977, *J. Chem. Phys*, 67, 3683-3704.

- Raina S et al., 1993, *J. Bacteriol.*, 175, 5009.
- Roux B et al., 1990, *J. Phys. Chem.*, 94, 4683-4688.
- Ryu S-E et al., 1990, *Nature*, 348, 419.
- Schmoll et al., 1990, *Microb. Pathog.*, 9, 331.
- 5 Simons B L et al., 1990, *FEMS Microbiol. Lett.*, 67, 107-112.
- Slonim L et al, 1992, *EMBO J.* 11, 4747-56.
- Strauch K L, Johnson K and Beckwith J., 1989. *J. Bacteriol.* 171, 2689-2696.
- Striker R et al., 1994, *J. Biol. Chem.*, 269, 4747-56.
- 10 Stromberg N et al., 1990, *EMBO J.*, 9, 2001-2101.
- Stromberg N et al., 1991, *Proc. Natl. Acad. Sci. USA*, 88, 9340-9344.
- Ullman J R, 1976, *J. Ass. Comp. Machinery*, 23, 31-42.
- de Vos A M et al, 1992, *Science*, 255, 306.
- 15 Wang J et al, 1990, *Nature*, 348, 411.
- Warshel A and Russell S T, 1984, *Q. Rev. Biophys.*, 17, 283-422.
- Willems R J L et al., 1992, *Molec. Microbiol.*, 6, 2661.
- Xuong N-H et al., 1985, *J. Appl. Crystallogr.*, 18, 342.
- 20 Åqvist J, 1990, *J. Phys. Chem.*, 94, 8021-8024.

167

Åqvist J and Medina C, 1993, 206th ACS National Meeting,  
Phys: 124, Catalogue of Abstracts.

Åqvist J et al., 1994, Protein Engineering, 7(3), 385-391.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

- (A) NAME: Washington University
- (B) STREET: 724 South Euclide Avenue
- (C) CITY: St. Louis
- (D) STATE: Missouri
- (E) COUNTRY: USA
- (F) POSTAL CODE: MO 63110

## (i) APPLICANT:

- (A) NAME: Symbicom Aktiebolag
- (B) STREET: Tvistevaegen 48
- (C) CITY: Umea
- (E) COUNTRY: Sweden
- (F) POSTAL CODE: S-907 36 Umea

(ii) TITLE OF INVENTION: TREATMENT OR PROPHYLAXIS OF DISEASES CAUSED BY PILUS-FORMING BACTERIA

(iii) NUMBER OF SEQUENCES: 27

## (iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 874 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Escherichia coli

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..720

## (ix) FEATURE:

- (A) NAME/KEY: sig\_peptide
- (B) LOCATION: 1..63

## (ix) FEATURE:

- (A) NAME/KEY: mat\_peptide

(B) LOCATION: 64..717

## (x) PUBLICATION INFORMATION:

(A) AUTHORS: Norgren, M.

Baga, M.

Tennet, J.M.

Normark, S.

(C) JOURNAL: Mol. Biol. Rep.

(D) VOLUME: 12

(F) PAGES: 169-178

(G) DATE: 1987

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATG ATT CGA AAA AAG ATT CTG ATG GCT GCC ATC CCC CTG TTT GTT ATA	48
Met Ile Arg Lys Lys Ile Leu Met Ala Ala Ile Pro Leu Phe Val Ile	
-21 -20 -15 -10	
TCC GGG GCA GAC GCT GCT GTT TCG CTG GAC AGA ACC CGC GCG GTG TTT	96
Ser Gly Ala Asp Ala Ala Val Ser Leu Asp Arg Thr Arg Ala Val Phe	
-5 1 5 10	
GAC GGG AGT GAG AAG TCA ATG ACG CTT GAT ATC TCC AAT GAT AAC AAA	144
Asp Gly Ser Glu Lys Ser Met Thr Leu Asp Ile Ser Asn Asp Asn Lys	
15 20 25	
CAA CTG CCC TAT CTT GCT CAG GCA TGG ATA GAA AAT GAA AAT CAG GAA	192
Gln Leu Pro Tyr Leu Ala Gln Ala Trp Ile Glu Asn Glu Asn Gln Glu	
30 35 40	
AAA ATT ATT ACA GGG CCG GTT ATT GCC ACC CCT CCG GTT CAG CGC CTT	240
Lys Ile Ile Thr Gly Pro Val Ile Ala Thr Pro Pro Val Gln Arg Leu	
45 50 55	
GAG CCG GGT GCG AAA AGC ATG GTC AGG CTG AGT ACC ACA CCG GAT ATC	288
Glu Pro Gly Ala Lys Ser Met Val Arg Leu Ser Thr Thr Pro Asp Ile	
60 65 70 75	
AGT AAA CTT CCT CAG GAC AGG GAA TCA CTG TTT TAT TTT AAT CTC AGG	336
Ser Lys Leu Pro Gln Asp Arg Glu Ser Leu Phe Tyr Phe Asn Leu Arg	
80 85 90	
GAA ATA CCG CCG AGG AGT GAA AAG GCC AAT GTA CTG CAG ATA GCC TTA	384
Glu Ile Pro Pro Arg Ser Glu Lys Ala Asn Val Leu Gln Ile Ala Leu	
95 100 105	
CAG ACC AAA ATA AAG CTT TTT TAT CGC CCG GCA GCA ATT AAA ACC AGA	432
Gln Thr Lys Ile Lys Leu Phe Tyr Arg Pro Ala Ala Ile Lys Thr Arg	
110 115 120	
CCA AAT GAA GTA TGG CAG GAC CAG TTA ATT CTG AAC AAA GTC AGC GGT	480
Pro Asn Glu Val Trp Gln Asp Gln Leu Ile Leu Asn Lys Val Ser Gly	
125 130 135	
GGG TAT CGT ATT GAA AAC CCA ACG CCC TAT TAT GTC ACT GTT ATT GGT	528
Gly Tyr Arg Ile Glu Asn Pro Thr Pro Tyr Tyr Val Thr Val Ile Gly	
140 145 150 155	

170

CTG GGA GGA AGT GAA AAG CAG GCA GAG GAA GGT GAG TTT GAA ACC GTG 576  
 Leu Gly Gly Ser Glu Lys Gln Ala Glu Glu Gly Glu Phe Glu Thr Val  
                     160                    165                    170

ATG CTG TCT CCC CGT TCA GAG CAG ACA GTA AAA TCG GCA AAT TAT AAT 624  
 Met Leu Ser Pro Arg Ser Glu Gln Thr Val Lys Ser Ala Asn Tyr Asn  
                     175                    180                    185

ACC CCT TAT CTG TCT TAT ATT AAT GAC TAT GGT GGT CGC CCG GTA CTG 672  
 Thr Pro Tyr Leu Ser Tyr Ile Asn Asp Tyr Gly Gly Arg Pro Val Leu  
                     190                    195                    200

TCG TTT ATC TGT AAT GGT AGC CGT TGC TCT GTG AAA AAA GAG AAA TAATGTACCG 727  
 Ser Phe Ile Cys Asn Gly Ser Arg Cys Ser Val Lys Lys Glu Lys  
                     205                    210                    215

CAATAACGGT TAAATGCGGG TGGGATATTA TGGTTGTGAA TAAAACAACA GCAGTACTGT 787

ATCTTATTGC ACTGTCGCTG AGTGGTTTCA TCCATACTTT CCTGCGGGCT GAAGAGCGGG 847

GTATATACGA TGACGTCTTT ACTGCAG 874

## (2) INFORMATION FOR SEQ ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 239 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Ile Arg Lys Lys Ile Leu Met Ala Ala Ile Pro Leu Phe Val Ile  
 -21 -20                    -15                    -10

Ser Gly Ala Asp Ala Ala Val Ser Leu Asp Arg Thr Arg Ala Val Phe  
 -5                    1                    5                    10

Asp Gly Ser Glu Lys Ser Met Thr Leu Asp Ile Ser Asn Asp Asn Lys  
                     15                    20                    25

Gln Leu Pro Tyr Leu Ala Gln Ala Trp Ile Glu Asn Glu Asn Gln Glu  
                     30                    35                    40

Lys Ile Ile Thr Gly Pro Val Ile Ala Thr Pro Pro Val Gln Arg Leu  
                     45                    50                    55

Glu Pro Gly Ala Lys Ser Met Val Arg Leu Ser Thr Thr Pro Asp Ile  
                     60                    65                    70                    75

Ser Lys Leu Pro Gln Asp Arg Glu Ser Leu Phe Tyr Phe Asn Leu Arg  
                     80                    85                    90

Glu Ile Pro Pro Arg Ser Glu Lys Ala Asn Val Leu Gln Ile Ala Leu  
                     95                    100                    105



171

Gln Thr Lys Ile Lys Leu Phe Tyr Arg Pro Ala Ala Ile Lys Thr Arg  
 110 115 120

Pro Asn Glu Val Trp Gln Asp Gln Leu Ile Leu Asn Lys Val Ser Gly  
 125 130 135

Gly Tyr Arg Ile Glu Asn Pro Thr Pro Tyr Tyr Val Thr Val Ile Gly  
 140 145 150 155

Leu Gly Gly Ser Glu Lys Gln Ala Glu Glu Gly Glu Phe Glu Thr Val  
 160 165 170

Met Leu Ser Pro Arg Ser Glu Gln Thr Val Lys Ser Ala Asn Tyr Asn  
 175 180 185

Thr Pro Tyr Leu Ser Tyr Ile Asn Asp Tyr Gly Gly Arg Pro Val Leu  
 190 195 200

Ser Phe Ile Cys Asn Gly Ser Arg Cys Ser Val Lys Lys Glu Lys  
 205 210 215

## (2) INFORMATION FOR SEQ ID NO: 3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (Synthetic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GTCAAACACC GCCGGAATC GTCCAGGCGA

30

## (2) INFORMATION FOR SEQ ID NO: 4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (Synthetic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CGGGCGATAA AAAAGAGCTA TTTTGGTCTG

30

## (2) INFORMATION FOR SEQ ID NO: 5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

172

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (Synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GCGATAAAAA AGCATTATTT TCCTCTG

27

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Gly	Lys	Arg	Lys	Pro	Gly	Glu	Leu	Ser	Gly	Ser	Met	Thr	Met	Val	Leu
1				5				10						15	

Ser Phe Pro

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Lys	Pro	Gly	Glu	Leu	Ser	Gly	Ser	Met	Thr	Met	Val	Leu	Ser	Phe	Pro
1				5				10					15		

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Ser	Gly	Ser	Met	Thr	Met	Val	Leu	Ser	Phe	Pro
1				5				10		

## (2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 7 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Thr Met Val Leu Ser Phe Pro  
1 5

## (2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Gly Lys Arg Lys Pro Val Glu Leu Ser Gly Ser Met Thr Met Val Leu  
1 5 10 15

Ser Ser Pro

## (2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Gly Lys Arg Lys Pro Gly Glu Leu Ser Gly Ser Met Thr Met Val Leu  
1 5 10 15

Ser Phe Pro

## (2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 amino acids  
(B) TYPE: amino acid

174

- (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Glu Glu Gly Lys Arg Lys Pro Gly Glu Leu Ser Gly Ser Met Thr Met  
1                      5                      10                      15

Val Leu Ser Phe  
20

(2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

Gln Asn Leu Ile Ala Gly Pro Phe Ser Ala Thr Ala Thr Leu Val Ala  
1                      5                      10                      15

Ser Tyr Ser

(2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Lys Lys Leu Glu Ala Gly Asn Tyr Phe Ala Val Leu Gly Phe Arg Val  
1                      5                      10                      15

Asp Tyr Glu

(2) INFORMATION FOR SEQ ID NO: 15:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

175

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

Lys Ser Val Val Pro Gly Asp Tyr Glu Ala Thr Ala Thr Phe Glu Leu  
1                      5                      10                      15  
Thr Tyr Arg

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Gly Ile Leu Asn Gly Gly Asp Phe Gln Thr Thr Ala Ser Met Ala Met  
1                      5                      10                      15  
Ile Tyr Asn

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

Tyr Ala Leu Ala Pro Asn Ala Val Ile Pro Thr Ser Leu Ala Leu Leu  
1                      5                      10                      15

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

176

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Lys Ser Val Val Pro Gly Asp Tyr Glu Ala Thr Ala Thr Phe Glu Leu  
1                      5                      10                      15

Thr Tyr Arg

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Leu Pro Ala Thr Asn Thr Leu Met Leu Ser Phe Asp Asn Val Gly Gly  
1                      5                      10                      15

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

Asp Gln Ile Lys Gln Leu Pro Ala Thr Asn Thr Leu Met Leu Ser Phe  
1                      5                      10                      15

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

177

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

Tyr	Lys	Met	Pro	Tyr	Asp	Gln	Ile	Lys	Gln	Leu	Pro	Ala	Thr	Asn	Thr
1				5					10					15	

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

Gln	His	His	Tyr	Tyr	Asp	Leu	Trp	Gln	Asp	His	Tyr	Lys	Met	Pro	Tyr
1				5					10					15	

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

CCCCCCTGCA GATCAGATTA AGCAGCTACC TGC

33

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

CCCCCTGCAGT AAAAATATCT CTGCTCAGAA ATAC

34

(2) INFORMATION FOR SEQ ID NO: 25:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 28 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

CCATCGATGA ACAGCCAGTC AGATAATC

28

(2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 61 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

GGAAAGAGAA AACCCGGGGA GCTATCTGGT TCTATGACTA TGGTTCTGAG TTTCCCTCTGA

60

T

61

(2) INFORMATION FOR SEQ ID NO: 27:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 67 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:



WO 95/14028

PCT/US94/13455

179

CGATCAGGGG AACTCAGAA CCATAGTCAT AGAACCAGAT AGCTCCCCGG GTTTTCTCTT

60

TCCTGCA

67

## CLAIMS

1. A method for the treatment and/or prophylaxis of diseases caused by tissue-adhering pilus-forming bacteria, comprising preventing, inhibiting or enhancing binding between at least one type of pilus subunit and at least one type of periplasmic molecular chaperone in the pilus-forming bacteria, the periplasmic molecular chaperone being one that binds pilus subunits during transport of these pilus subunits through the periplasmic space and/or during the process of assembly of the intact pilus.
2. A method according to claim 1, comprising administering, to a subject in need thereof, an effective amount of a substance capable of interacting with at least one type of periplasmic molecular chaperone which binds pilus subunits during transport of these pilus subunits through the periplasmic space and/or during the process of assembly of the intact pilus, in such a manner that binding of pilus subunits to the periplasmic molecular chaperone is prevented, inhibited or enhanced.
3. A method according to claim 1 or 2, wherein the bacteria are selected from the group consisting of *Haemophilus* spp, *Helicobacter* spp, *Pseudomonas aeruginosa*, *Mycoplasma* spp, and all members of the *Enterobacteriaceae* family, including *Escherichia* spp, *Salmonella* spp, *Bordetella* spp, *Yersinia* spp, *Klebsiella* spp., and *Proteus* spp.
4. A method according to any of claims 1-3, wherein the periplasmic molecular chaperone is a periplasmic protein selected from the group consisting of *PapD*, *FimC*, *SfaE*, *FaeE*, *FanE*, *Cs3-1*, *F17D*, *ClpE*, *EcpD*, *Mrkb*, *FimB*, *SefB*, *HifB*, *MyfB*, *PsaB*, *PefD*, *YehC*, *MrpD*, *CssC*, *NfaE*, *AggD*, and *Caf1M*.
5. A method according to any of claims 1-4, wherein the prevention, inhibition or enhancement of the binding is accomplished by interaction with, in the periplasmic molecu-

lar chaperone, a binding site which is normally involved in binding to pilus subunits during transport of these pilus subunits through the periplasmic space and/or during the process of pilus assembly.

5 6. A method according to claim 5, wherein the binding site is a binding site to which the carboxyl terminal part of a pilus subunit is capable of binding, and which comprises site points substantially identical to the invariant residues Arg-8 and Lys-112 in *PapD*, and a polypeptide fragment which  
10 is capable of interacting with a  $\beta$ -strand of the carboxyl terminal part of the pilus subunit, thereby stabilizing the binding of said subunit at the Arg-8 and Lys-112 site points of the binding site.

7. A method according to claim 5, wherein the binding site is  
15 one to which the peptide G125'-140' and/or the fusion peptide MBP-G1'-140' and/or the peptide G1'-19' is capable of binding.

8. A method for identifying a potentially therapeutically useful substance capable of interacting with a periplasmic molecular chaperone, thereby preventing, inhibiting or enhan-  
20 cing the interaction between a periplasmic molecular chaperone and a pilus subunit, the method comprising at least one of the following steps:

1) testing a candidate substance in an assay in which the possible prevention, inhibition or enhancement by the sub-  
25 stance of the interaction between the periplasmic molecular chaperone and the pilus subunit is determined by

adding the substance to a system comprising the periplasmic molecular chaperone or an analogue thereof in an immobi-  
lized form and the pilus subunit or an equivalent thereof  
30 in a solubilized form and determining the change in binding between the pilus subunit or equivalent thereof and the periplasmic molecular chaperone or analogue thereof caused by the addition of the substance, or

adding the substance to a system comprising the pilus subunit or an equivalent thereof in an immobilized form and the periplasmic molecular chaperone or an analogue thereof in a solubilized form and determining the change in binding between the pilus subunit or equivalent thereof and the periplasmic molecular chaperone or analogue thereof caused by the addition of the substance, or

adding the substance to a system comprising the pilus subunit or an equivalent thereof as well as the periplasmic molecular chaperone or an analogue thereof in solubilized form and determining the change in binding between the pilus subunit or equivalent thereof and the periplasmic molecular chaperone or analogue thereof caused by the addition of the substance, or

adding the substance to a system comprising the pilus subunit or an equivalent thereof as well as the periplasmic molecular chaperone or an analogue thereof in solubilized form and measuring the change in binding energy caused by the addition of the substance, and identifying the substance as potentially therapeutically useful if a significant change in the binding energy between the pilus subunit or equivalent thereof and the periplasmic molecular chaperone or analogue thereof is observed,

and identifying the substance as potentially therapeutically useful if a significant change in the binding or binding energy between the pilus subunit or equivalent thereof and the periplasmic molecular chaperone or analogue thereof is observed;

2) testing a candidate substance in an assay in which the possible prevention, inhibition or enhancement of the interaction between the periplasmic molecular chaperone and the pilus subunit is determined by

adding the substance to a system comprising living tissue-  
adhering pilus-forming bacteria followed by determination  
of the growth rate of the bacteria, a reduction in growth  
rate compared to a corresponding system wherein the sub-  
stance has not been added being indicative of prevention,  
inhibition or enhancement of the binding between the peri-  
plasmic molecular chaperone and the pilus subunit, or

adding the substance to a system comprising living tissue-  
adhering pilus-forming bacteria followed by a determination  
of the tissue adhesion of the bacteria, a reduction in  
tissue adhesion compared to a corresponding system wherein  
the substance has not been added being indicative of pre-  
vention, inhibition or enhancement of the binding between  
the periplasmic molecular chaperone and the pilus subunit,

and identifying the substance as potentially therapeutically  
useful if a reduction in growth rate or tissue adhesion is  
observed after the addition of the substance; and

3) administering, to an experimental animal, a substance  
which has been established in vitro to prevent, inhibit  
or enhance the interaction between a periplasmic mole-  
cular chaperone and a pilus subunit, the experimental  
animal being inoculated with tissue-adhering pilus-  
forming bacteria before, simultaneously with or after  
the administration of the substance, and electing as a  
substance suitably capable of interacting with a peri-  
plasmic molecular chaperone, a substance preventing  
and/or curing and/or alleviate disease caused by the  
bacteria.

9. A method according to claim 8, wherein the assay in step 1  
comprises

- adding the substance to a first system comprising the  
periplasmic molecular chaperone or an analogue thereof,

- subsequently adding a pilus subunit or an equivalent thereof of which has been labelled with an environmentally sensitive fluorescent probe,
  - determining the fluorescent emission at a particular wavelength which is indicative of the amount of binding between the periplasmic molecular chaperone or the analogue thereof and the pilus subunit or the equivalent thereof, and
  - comparing the determined fluorescent emission to fluorescent emission determined in a corresponding second system containing substantially the same concentrations of the molecular chaperone or the analogue thereof and the pilus subunit or the equivalent thereof but substantially no substance,
- a significant difference in fluorescent emission between the first and second system being indicative of interaction between the periplasmic molecular chaperone or the analogue thereof and the substance.

10. A method according to claim 9, wherein the determination of fluorescent emission in the second system is performed a plurality of times at varying molar ratios between the pilus subunit or the equivalent thereof and the periplasmic chaperone and the equivalent thereof, whereupon the constant of binding between the pilus subunit or equivalent thereof and the periplasmic molecular chaperone or analogue thereof is assessed from the determined fluorescent emission data.

11. A method according to claim 9 or 10, wherein the determination of fluorescent emission in the first system is performed a plurality of times at varying molar ratios between the substance and the periplasmic chaperone and the equivalent thereof, whereafter the constant of binding between the substance and the periplasmic molecular chaperone or analogue thereof is assessed from the determined fluorescent emission data.

12. A method according to any of claims 8-11, wherein the equivalent of the pilus subunit is G1'-19'WT, MSP-G1'-140', or G125'-140' and the periplasmic molecular chaperone is PapD.
- 5 13. A method according to any of claims 8-12, wherein the living, tissue adhering pilus-forming bacteria are of a protease deficient strain, the protease being one which is at least partially responsible for the degradation of pilus subunits
- 10 14. A method according to claim 13, wherein the protease deficient strain is a *degP41* strain.
- 15 15. A method for identifying and/or designing a substance, X, capable of interacting with a periplasmic molecular chaperone, e.g. binding to the periplasmic molecular chaperone, with a predicted binding free energy equal to or better than a predetermined threshold value, the method comprising
- 20 1) selecting a substance, A, which could potentially interact with a site in the periplasmic molecular chaperone, and providing a 3-dimensional structural representation thereof,
- 2) predicting the binding free energy between the substance A and the site in the periplasmic molecular chaperone or a site analogous to such a site,
- 25 3) if the predicted binding free energy between the substance A and the site in the periplasmic molecular chaperone or the analogous site is equal to or better than the predetermined threshold value, then identifying the substance A as the substance X,
- 30 4) if the predicted binding free energy between the substance A and the site in the periplasmic molecular chaperone or the analogous site is not equal to or better than

a predetermined threshold value, then modifying the 3-dimensional structural representation and predicting the binding free energy between the thus modified substance, B, and the site in the periplasmic molecular chaperone or the analogous site, and

- 5
- 5) repeating step 4 until the predicted binding free energy determined between the resulting substance, X, and the site in the periplasmic molecular chaperone or the analogous site is equal to or better than the predetermined threshold value.
- 10

16. A method for identifying and/or designing a substance, Y, capable of interacting with a periplasmic molecular chaperone, e.g. binding to the periplasmic molecular chaperone, with a binding free energy equal to or better than a predetermined threshold value, the method comprising

15

- 1) Identifying a substance, X, according to the method of claim 15, and
- 2) providing a sample of the chemical substance X and a sample of the periplasmic molecular chaperone or an analogue thereof and measuring the binding free energy between the chemical substance X and the periplasmic molecular chaperone or analogue thereof, and establishing that the measured binding free energy between the chemical substance X and the periplasmic molecular chaperone or analogue thereof is equal to or better than the predetermined threshold value, and then identifying the substance X as the substance Y, and optionally
- 20
- 25
- 3) subjecting the substance Y to the method according to any of claims 8-14 to verify that the substance Y is a potentially therapeutically useful substance capable of interacting with a periplasmic molecular chaperone.
- 30



17. A method according to claim 15 or 16, wherein the binding free energy is predicted by

assessing the average energy difference,  $\langle \Delta V_{X-s}^{el} \rangle$ , defined as  $\langle V_{X-s}^{el} \rangle_B - \langle V_{X-s}^{el} \rangle_A$ , between the contribution from polar interactions to the potential energy between the chemical substance X and its surroundings (denoted s) in two states, one state (A) being where the chemical substance is surrounded by solvent, the other state (B) being where the chemical substance, bound to the site in the periplasmic molecular chaperone or bound to a site analogous to such a site, is surrounded by solvent,

assessing the average energy difference,  $\langle \Delta V_{X-s}^{vdw} \rangle$ , defined as  $\langle V_{X-s}^{vdw} \rangle_B - \langle V_{X-s}^{vdw} \rangle_A$ , between the contribution from non-polar interactions to the potential energy between the chemical substance X and its surroundings (denoted s) in two states, one state (A) being where the chemical substance is surrounded by solvent, the other state (B) being where the chemical substance, bound to the site in the periplasmic molecular chaperone or bound to a site analogous to such a site, is surrounded by solvent, and

calculating the absolute binding free energy as an adjusted combination of the two above-mentioned average energy differences.

18. The method according to claim 17, wherein the site in the periplasmic molecular chaperone is the pilus subunit binding part of a periplasmic molecular chaperone, such as the pilus subunit binding site of a periplasmic molecular chaperone selected from the group consisting of *PapD*, *FimC*, *SfaE*, *FaeE*, *FanE*, *Cs3-1*, *F17D*, *ClpE*, *EcpD*, *Mrkb*, *FimB*, *SefB*, *HifB*, *MyfB*, *PsaB*, *PefD*, *YehC*, *MrpD*, *CssC*, *NfaE*, *AggD* and *Caf1M*, or an analogue of such a pilus subunit binding site.

19. The method according to claim 18, wherein the site in the periplasmic molecular chaperone is a pilus subunit binding site of *PapD* or an analogue thereof.

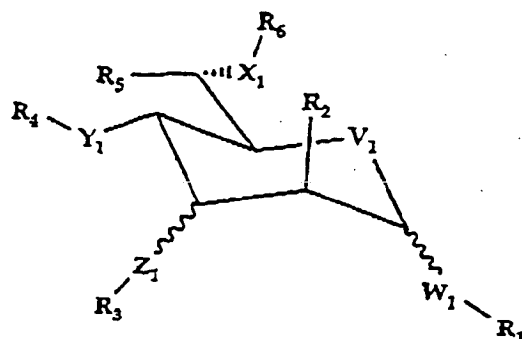
20. The method according to any of claims 15-19, wherein the substance A is chosen by performing the following steps:

- 10       - co-crystallizing the periplasmic molecular chaperone or the analogue thereof with a ligand capable of interacting with a site in the periplasmic molecular chaperone or the analogue thereof and establishing the three-dimensional conformation of the periplasmic molecular chaperone or the analogue thereof and the ligand when interacting by means of X-ray crystallography,
- 15       - using the above-established conformation of the periplasmic molecular chaperone or the analogue thereof to establish a 3-dimensional representation of the site in the periplasmic molecular chaperone or the analogue thereof interacting with the ligand during binding,
- 20       - selecting a number of distinct chemical groups, X1, and determining the possible spatial distributions of the X1 chemical groups which maximizes the binding free energy between the chemical groups and the site in the chaperone or the analogue interacting with the ligand,
- 25       - extracting, from a database comprising three-dimensional representations of molecules, a molecule which has the X1 chemical groups in the possible spatial distributions determined above,
- optionally modifying the 3-dimensional representation of the molecule extracted from the database, and

30       identifying the optionally modified molecule as the substance A.

21. A method according to claim 20, wherein the ligand is a pilus subunit or a part thereof with which the periplasmic molecular chaperone normally interacts during transport of the pilus subunit through the periplasmic space and/or during pilus assembly.

22. A novel compound of the general formula:



I

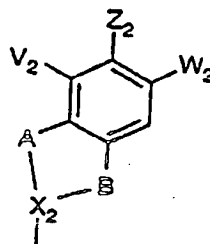
wherein

$V_1$  is O, S, SO,  $SO_2$ ,  $CH_2$ ,  $C(OH)H$ , CO or CS;

10  $W_1$  is O, S,  $SO_2$ ,  $SO_3$ ,  $CH_2$  or NH;

$R_1$  is H;  $C_{1-24}$  alkyl,  $C_{1-24}$  alkenyl or  $C_{1-24}$  alkynyl, which alkyl, alkenyl and alkynyl may be substituted with one or more substituents independently selected from OH,  $-CONH_2$ ,  $-CSNH_2$ ,  $-CONHOH$ ,  $-CSNHOH$ ,  $-NHCHO$ ,  $-NHCONH_2$ ,  $-NHCSNH_2$ ,  $-NHSO_2NH_2$  and  $-SO_2NH_2$ ; acyl; or  $-(CH_2CH_2O)_s-H$ , wherein  $s = 1, 2, 3$ ;

$R_2$  is a group of the formula



190

wherein

A is  $-\text{CH}-(\text{CH}_2)_n-$ , or  $-\text{CH}=\text{CH}-(\text{CH}_2)_{n-1}-$  ( $n>0$ ) or  $-\text{CH}-(\text{CH}_2)_n-$ ;

||

5

 $\text{Y}_2$ 

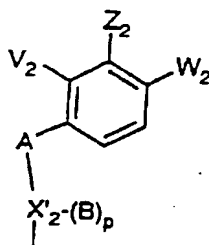
B is  $-(\text{CH}_2)_m-$  or  $=\text{CH}-(\text{CH}_2)_{m-1}-$  ( $m>0$ );

$\text{X}_2$  is N, CH or C (when B is  $=\text{CH}-(\text{CH}_2)_{m-1}-$ ; and

$\text{Y}_2$  is O, S, NH,  $\text{H}_2$  or H ( $n=1$ ); and

$4>m+n>0$ ,  $n<3$ , and  $m<3$ ;

10 or  $\text{R}_2$  is a group of the formula



wherein

A is  $-\text{CH}-(\text{CH}_2)_n-$ , or  $-\text{CH}=\text{CH}-(\text{CH}_2)_{n-1}-$  ( $n>0$ ) or  $-\text{CH}-(\text{CH}_2)_n-$ ;

||

15

 $\text{Y}_2$ 

B is  $-(\text{CH}_2)_m-$  or  $=\text{CH}-(\text{CH}_2)_{m-1}-$  ( $m>0$ ); and

$\text{X}'_2$  is O, NH,  $\text{CH}_2$  or S (when  $p=0$ ); N or CH ( $p=1$ ); or C (when  $p=1$  and B is  $=\text{CH}-(\text{CH}_2)_{m-1}-$ );

20

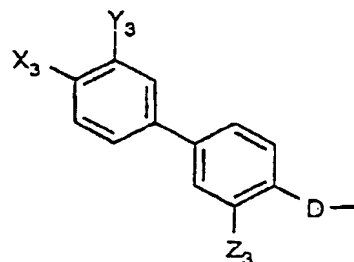
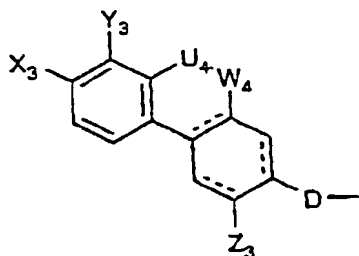
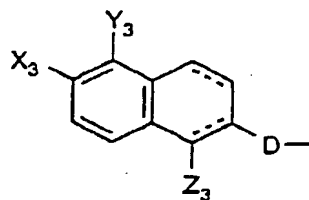
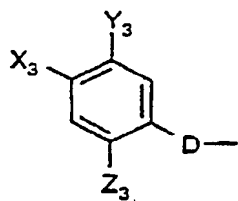
$\text{V}_2$ ,  $\text{Z}_2$  and  $\text{W}_2$  are independently H, OH,  $-\text{CONH}_2$ ,  $-\text{CSNH}_2$ ,  $-\text{CONHOH}$ ,  $-\text{CSNHOH}$ ,  $-\text{NHCHO}$ ,  $-\text{NHCONH}_2$ ,  $-\text{NHCSNH}_2$ ,  $-\text{NHSO}_2\text{NH}_2$ ,  $-\text{SO}_2\text{NH}_2$ , or  $\text{V}_2$  and  $\text{Z}_2$ , or  $\text{Z}_2$  and  $\text{W}_2$  together form  $-\text{NHC}(\text{O})\text{NH}-$ ,  $-\text{C}(\text{O})\text{NHC}(\text{O})-$ ,  $-\text{NHS}(\text{O}_2)\text{NH}-$ ,  $-\text{C}(\text{O})\text{NHO}-$ ,  $-\text{C}(\text{S})\text{NHO}-$ ,  $-\text{S}(\text{O}_2)\text{NHO}-$ , or  $-\text{S}(\text{O}_2)\text{NHC}(\text{O})-$ ;

$4>m+n>0$ ,  $n<3$ , and  $m<3$ ;

or  $R_2$  is a group  $-W_5-(C_{1-5} \text{ alkyl or } C_{2-5} \text{ alkenyl or } C_{2-5} \text{ alkynyl})$  wherein  $W_5$  is a bond or is selected from  $-O-$ ,  $-S-$ ,  $-SO_2-$ , and  $-NHC(O)-$ , and the  $C_{1-5}$  alkyl,  $C_{2-5}$  alkenyl or  $C_{2-5}$  alkynyl moiety may be substituted with up to three groups  
 5 selected independently from  $OH$ ,  $-CONH_2$ ,  $-CSNH_2$ ,  $-CONHOH$ ,  $-CSNHOH$ ,  $-NHCHO$ ,  $-NHCONH_2$ ,  $-NHCSNH_2$ ,  $-NHSO_2NH_2$  and  $-SO_2NH_2$ ;

$-Z_1-R_3$  is  $-SO_2(OH)$ ,  $-PO(OH)_2$ ,  $-OSO_2(OH)$ ,  $-NHSO_2(OH)$ ,  $-SPO(OH)_2$ ,  $-CH_2COOH$ , tetrazol-5-yl or tetrazol-5-ylmethyl, or salts thereof;

10 or  $Z_1$  is  $-O-$ ,  $-S-$ ,  $-NH-$ , or  $-CH_2-$ , and  $R_3$  is a group of the formula:



$D$  is  $-CH_2-$ ,  $-CO-$ ,  $-SO_2-$ ,  $-NH-SO_2-$ ,  $-NH-CO-$ ,  $-O-PO(OH)-$  or a salt thereof;

15  $Z_3$  is  $H$ ,  $OH$ ,  $-CONH_2$ ,  $-CSNH_2$ ,  $-CONHOH$ ,  $-CSNHOH$ ,  $-NHCHO$ ,  $-NHCONH_2$ ,  $-NHCSNH_2$ ,  $-NHSO_2NH_2$ ,  $-SO_2NH_2$ ,  $-SO_2(OH)$ ,  $-PO(OH)_2$ ,  $-OSO_2(OH)$ ,  $-NHSO_2(OH)$ ,  $-COOH$ , tetrazolyl-5-yl or tetrazolyl-5-ylmethyl or a salt thereof,

with the proviso that when D is  $-\text{CH}_2-$ ,  $-\text{CO}-$ ,  $-\text{SO}_2-$ ,  $-\text{NHSO}_2-$  or  $-\text{NHCO}-$ , then  $\text{Z}_3$  is  $-\text{SO}_2(\text{OH})$ ,  $-\text{PO}(\text{OH})_2$ ,  $-\text{OSO}_2(\text{OH})$ ,  $-\text{NHSO}_2(\text{OH})$ ,  $-\text{COOH}$ , tetrazolyl-5-yl or tetrazolyl-5-ylmethyl or a salt thereof;

- 5  $\text{X}_3$  and  $\text{Y}_3$  independently are H,  $\text{NO}_2$ ,  $\text{SO}_2\text{NH}_2$ ,  $\text{CONH}_2$ ,  $\text{CF}_3$  or F; and

$\text{U}_4-\text{W}_4$  is  $-\text{CHCH}-$ ,  $-\text{CH}_2\text{CH}_2-$ ,  $-\text{C}(\text{OH})\text{CH}-$ ,  $-\text{CHC}(\text{OH})-$ ,  $-\text{CH}(\text{OH})\text{CH}_2-$ ,  $-\text{CH}_2\text{CH}(\text{OH})-$ ,  $-\text{CH}(\text{OH})\text{CH}(\text{OH})-$ ,  $-\text{C}(\text{O})\text{NH}-$ ,  $-\text{NHC}(\text{O})-$ ;

- 10  $\text{Y}_1$  is  $-\text{O}-$  or  $-\text{S}-$ ;

$\text{R}_4$  is H or, when  $\text{Y}_1$  is S,  $\text{S}(\text{CH}_2)_q\text{N}(\text{R}_9)_3^+$  and q is an integer 2-4, where  $\text{R}_9$  is H or  $\text{CH}_3$ ;

- $\text{R}_5$  is H;  $\text{C}_{1-6}$  alkyl,  $\text{C}_{2-6}$  alkenyl or  $\text{C}_{2-6}$  alkynyl, and the  $\text{C}_{1-6}$  alkyl,  $\text{C}_{2-6}$  alkenyl or  $\text{C}_{2-6}$  alkynyl moiety may be substituted  
 15 with OH,  $-\text{CONH}_2$ ,  $-\text{CSNH}_2$ ,  $-\text{CONHOH}$ ,  $-\text{CSNHOH}$ ,  $-\text{NHCHO}$ ,  $-\text{NHCONH}_2$ ,  $-\text{NHCSNH}_2$ ,  $-\text{NHSO}_2\text{NH}_2$  or  $-\text{SO}_2\text{NH}_2$ ; or aryl, aryl( $\text{C}_{1-2}$ )alkyl, heterocyclyl, or heterocyclyl( $\text{C}_{1-2}$ )alkyl which may optionally be substituted in the aryl or heterocyclyl moieties with one,  
 20 two or three substituents selected independently from OH, F, Cl,  $\text{NH}_2$ ,  $\text{CONH}_2$ ,  $\text{NHCOH}$ , and  $\text{SO}_2\text{NH}_2$ ;

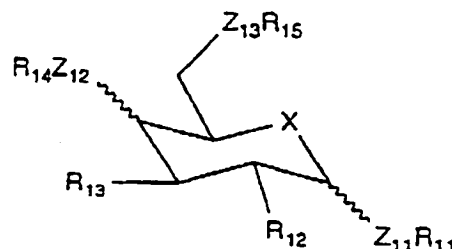
$\text{X}_1$  is  $-\text{O}-$ ,  $-\text{S}-$  or  $-\text{NH}-$ ;

$\text{R}_6$  is H or, when  $\text{X}_1$  is NH, acyl,  $\text{HOCNH-Val-Met-}$ ,  $\text{HOCNH-Ile-}$ ,  $(\text{S,S})$ -dioxo-methionyl- or  $\text{HOCNH-Val-(pyran-4-on-2-yl)-ala-}$ nyl-;

- 25 or a salt thereof.

23. A novel compound of the general formula

193



II

wherein

X is O, P, P(O), S, SO, SO<sub>2</sub>, CH<sub>2</sub>, C(OH)H, or a group NQ<sub>11</sub>, wherein Q<sub>11</sub> is H, OH, C<sub>1-24</sub> acyl or C<sub>1-24</sub> alkyl;

5 Z<sub>11</sub> is a bond, O, CH<sub>2</sub>, S, SO, SO<sub>2</sub>, or a group NQ<sub>12</sub>, wherein Q<sub>12</sub> is H, C<sub>1-24</sub> acyl or C<sub>1-24</sub> alkyl;

R<sub>11</sub> is H; C<sub>1-24</sub> alkyl, C<sub>1-24</sub> alkenyl, or C<sub>1-24</sub> alkynyl, which may be substituted with one or more substituents independently selected from -OH, -COOH, -F, -Cl, -CONH<sub>2</sub>, -CSNH<sub>2</sub>,  
 10 -CONHOH, -CSNHOH, -NHCHO, -NHCONH<sub>2</sub>, -NHCSNH<sub>2</sub>, -NHSO<sub>3</sub>NH<sub>2</sub> and -SO<sub>2</sub>NH<sub>2</sub>; acyl; or -(CH<sub>2</sub>CH<sub>2</sub>O)<sub>s</sub>-H, wherein s = 1, 2, or 3;

or R<sub>11</sub> is CH=CH-(CH<sub>2</sub>)<sub>n'</sub>-Q<sub>13</sub> or -(CH<sub>2</sub>)<sub>n'</sub>-Q<sub>13</sub>, wherein Q<sub>13</sub> is an aryl or a heteroaryl group substituted with OH, COOH, F, Cl, -CONH<sub>2</sub>, -CSNH<sub>2</sub>, -CONHOH, -CSNHOH, -NHCHO, -NHCONH<sub>2</sub>, -NHCSNH<sub>2</sub>,  
 15 -NHSO<sub>3</sub>NH<sub>2</sub> and -SO<sub>2</sub>NH<sub>2</sub>, and wherein n' ≥ 0;

R<sub>12</sub> and R<sub>13</sub> are independently OH, H, F, Cl, OW<sub>11</sub>, or O(CO)W<sub>11</sub>, wherein W<sub>11</sub> is C<sub>1-24</sub> alkyl, C<sub>1-24</sub> alkenyl or C<sub>1-24</sub> alkynyl, or an aryl or a heteroaryl group substituted with -OH, -COOH, -F, -Cl, -CONH<sub>2</sub>, -CSNH<sub>2</sub>, -CONHOH, -CSNHOH, -NHCHO, -NHCONH<sub>2</sub>,  
 20 -NHCSNH<sub>2</sub>, -NHSO<sub>3</sub>NH<sub>2</sub> and -SO<sub>2</sub>NH<sub>2</sub>;

Z<sub>12</sub> is a bond, O, S, or CH<sub>2</sub>;

R<sub>14</sub> is -(CH<sub>2</sub>)<sub>n''</sub>-Q<sub>14</sub>, wherein Q<sub>14</sub> is an aryl group or a heteroaryl group substituted with -OH, -COOH, -F, -Cl, -CONH<sub>2</sub>, -CSNH<sub>2</sub>, -CONHOH, -CSNHOH, -NHCHO, -NHCONH<sub>2</sub>, -NHCSNH<sub>2</sub>,  
 25 -NHSO<sub>3</sub>NH<sub>2</sub> and -SO<sub>2</sub>NH<sub>2</sub>, and wherein n'' = 0, 1, 2, or 3;

$Z_{13}$  is a bond, O,  $CH_2$ , S, SO,  $SO_2$ ,  $NQ_{14}Q_{15}$ , wherein  $Q_{14}$  is H,  $C_{1-24}$  acyl or  $C_{1-24}$  alkyl, and  $Q_{15}$  is CO or  $-C(O)W_{12}$ , wherein  $W_{12}$  is O or  $NW_{13}$ , wherein  $W_{13}$  is H, OH,  $C_{1-24}$  acyl or  $C_{1-24}$  alkyl;

- 5  $R_{15}$  is H;  $C_{1-24}$  alkyl,  $C_{1-24}$  alkenyl or  $C_{1-24}$  alkynyl, which alkyl, alkenyl and alkynyl may be substituted with one or more substituents independently selected from, -OH, -COOH, -F, -Cl, -CONH<sub>2</sub>, -CSNH<sub>2</sub>, -CONHOH, -CSNHOH, -NHCHO, -NHCONH<sub>2</sub>, -NHCSNH<sub>2</sub>, -NHSO<sub>3</sub>NH<sub>2</sub> and -SO<sub>2</sub>NH<sub>2</sub>; acyl or  $-(CH_2CH_2O)_s-H$ , where-
- 10 in  $s = 1, 2, \text{ or } 3$ ;

or  $R_{15}$  is  $CH=CH-(CH_2)_{n'}-Q_{13}$ , or  $-(CH_2)_{n'}-Q_{13}$ , wherein  $Q_3$  is as defined above and wherein  $n' \geq 0$ ;

or a salt thereof.

- 15 24. A compound according to claim 22 or 23, which causes a prevention, inhibition or enhancement of the binding of G1'-19'WT to *PapD*, and/or causes a prevention, inhibition or enhancement of the binding of MBP-G140' to *PapD*, and/or causes a prevention, inhibition or enhancement of the binding
- 20 of G125'-140' to *PapD* and/or causes a prevention, inhibition or enhancement of a pilus subunit to *PapD*.

25. A novel pyranoside or a salt thereof, which is selected from the group consisting of

- 25 Ethyl 2,3-O-Dibenzoyl-4-O-benzyl-1-thio- $\beta$ -D-glucohexopyranoside;
- Ethyl 6-O-acetyl-2,3-O-dibenzoyl-4-O-benzyl-1-thio- $\beta$ -D-glucohexopyranoside;
- Methylglycolyl 6-O-acetyl-2,3-O-dibenzoyl-4-O-benzyl- $\beta$ -D-glucohexopyranoside;
- 30 2-(Hydroxy)ethyl 4-O-benzyl- $\beta$ -D-glucopyranoside;
- Sodium glycolyl 4-O-benzyl- $\beta$ -D-glucohexopyranoside;
- Methyl 2-O-ethyl-4,6-O-(4'-methoxy)phenylmethylene- $\alpha$ -D-mannohexopyranoside;



- Methyl 2-O-ethyl-3-O-dimethyl-t-butylsilyl-4,6-O-(4'-methoxy)phenylmethylen- $\alpha$ -D-mannohexopyranoside;  
Methyl 2-O-ethyl-3-O-dimethyl-t-butylsilyl-4,-O-(4'-methoxy)-benzyl- $\alpha$ -D-mannohexopyranoside;  
5 methyl 2-O-ethyl-3-O-dimethyl-t-butylsilyl-6-O-(4'-methoxy)-benzyl- $\alpha$ -D-mannohexopyranoside;  
Methyl 2-O-ethyl-3-O-dimethyl-t-butylsilyl-4,-O-(4'-methoxy)-benzyl-6(S)-phenyl- $\alpha$ -D-mannohexopyranoside;  
Methyl 2,3-anhydro-4,6-O-p-methoxybenzylidene- $\alpha$ -D-mannopyranoside;  
10 Methyl 3-azido-4,6-O-p-methoxybenzylidene- $\alpha$ -D-altropyranoside;  
Methyl 3-azido-2-O-ethyl-4,6-O-p-methoxybenzylidene- $\alpha$ -D-altropyranoside;  
15 Methyl 3-azido-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl- $\alpha$ -D-altropyranoside;  
Methyl 3-azido-6-O-benzoyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl- $\alpha$ -D-altropyranoside;  
Methyl 6-O-benzoyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl-3-sulfamino- $\alpha$ -D-altropyranoside sodium salt;  
20 Methyl 6-O-benzoyl-3-deoxy-2-O-ethyl-3-sulfamino- $\alpha$ -D-altropyranoside ammonium salt;  
Methyl 3-azido-6-O-pivaloyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl- $\alpha$ -D-altropyranoside;  
25 Methyl 6-O-pivaloyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl-3-sulfamino- $\alpha$ -D-altropyranoside sodium salt;  
Methyl 6-O-pivaloyl-3-deoxy-2-O-ethyl-3-sulfamino- $\alpha$ -D-altropyranoside ammonium salt;  
Methyl 6-O-pivaloyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl-3-tbutyloxamido- $\alpha$ -D-altropyranoside;  
30 Methyl 6-O-pivaloyl-3-deoxy-2-O-ethyl-3-oxamido- $\alpha$ -D-altropyranoside ammonium salt;  
Methyl 3-azido-6-O-pyrrol-3'-ylcarboxyl-3-deoxy-2-O-ethyl-4-O-p-methoxybenzyl- $\alpha$ -D-altropyranoside; and  
35 Methyl 6-O-pyrrol-3'-ylcarboxyl-3-deoxy-2-O-ethyl-3-sulfamino- $\alpha$ -D-altropyranoside ammonium salt.

26. A method for identifying a binding site in a periplasmic molecular chaperone or an analogue thereof, comprising

co-crystallizing the periplasmic molecular chaperone or an analogue thereof with a ligand binding to the periplasmic molecular chaperone or the analogue thereof,

resolving the three-dimensional structure of the chaperone/ligand interaction, thereby resolving the three-dimensional structure of the periplasmic molecular chaperone or the analogue thereof when binding to the ligand,

determining the site-point(s) involved in the intermolecular interaction between the periplasmic molecular chaperone or the analogue thereof and the ligand, and

identifying the thus determined site-point(s) of the periplasmic molecular chaperone or the analogue thereof as a binding site in the periplasmic molecular chaperone or the analogue thereof.

27. A pharmaceutical composition, comprising, as an active compound, a substance capable of interacting with at least one type of periplasmic molecular chaperone which binds pilus subunits during transport of these pilus subunits through the periplasmic space and/or during the process of assembly of the intact pilus, in such a manner that binding of pilus subunits to the periplasmic molecular chaperone is prevented, inhibited or enhanced, in combination with at least one pharmaceutically acceptable carrier or excipient.

28. A pharmaceutical composition, comprising, as an active compound, a substance according to any of claims 22-25 or a substance identified according to the method in any of claims 8-21, in combination with at least one pharmaceutically acceptable carrier or excipient.

29. A pharmaceutical composition according to claim 28, which further comprises at least one additional pharmaceutical substance.

30. A pharmaceutical composition according to claim 29,  
5 wherein the at least one additional pharmaceutical substance is an antibacterial agent selected from the group consisting of penicillins, cephalosporins, aminoglycosides, sulfonamides, tetracyclines, chloramphenicol, polymyxins, antimycobacterial drugs and urinary antiseptics.

10 31. A substance for use as a pharmaceutical, the substance being capable of interacting with at least one type of periplasmic molecular chaperone which binds pilus subunits during transport of these pilus subunits through the periplasmic space and/or during the process of assembly of the intact  
15 pilus, in such a manner that binding of pilus subunits to the periplasmic molecular chaperone is prevented, inhibited or enhanced.

32. A substance according to claim 31, wherein the binding site is a binding site to which the carboxyl terminal part of  
20 a pilus subunit is capable of binding, and which comprises site points substantially identical to the invariant residues Arg-8 and Lys-112 in PapD, and a polypeptide fragment which is capable of interacting with a  $\beta$ -strand of the carboxyl terminal part of the pilus subunit, thereby stabilizing the  
25 binding of said subunit at the Arg-8 and Lys-112 site points of the binding site.

33. A substance according to claim 32 which causes a prevention, inhibition or enhancement of the binding of G1'-19'WT to PapD, and/or causes a prevention, inhibition or enhancement of the binding of MBP-G140' to PapD, and/or causes a  
30 prevention, inhibition or enhancement of the binding of G125'-140' to PapD and/or causes a prevention, inhibition or enhancement of a pilus subunit to PapD.

34. A substance according to any of claims 31-33 which has been identified according to the method as claimed in any of claims 8-21 or which is a compound according to any of claims 22-25.

5 35. The use of a substance which is capable of interacting with at least one type of periplasmic molecular chaperone which binds pilus subunits during transport of these pilus subunits through the periplasmic space and/or during the process of assembly of the intact pilus, in such a manner  
10 that binding of pilus subunits to the periplasmic molecular chaperone is prevented, inhibited or enhanced, for the preparation of a pharmaceutical composition for the treatment and/or prophylaxis of bacterial infection.

15 36. The use according to claim 35, wherein the binding site is a binding site to which the carboxyl terminal part of a pilus subunit is capable of binding, and which comprises site points substantially identical to the invariant residues Arg-8 and Lys-112 in *PapD*, and a polypeptide fragment which is capable of interacting with a  $\beta$ -strand of the carboxyl  
20 terminal part of the pilus subunit, thereby stabilizing the binding of said subunit at the Arg-8 and Lys-112 site points of the binding site.

25 37. The use according to claim 35 or 36, wherein the substance has been identified according to the method as claimed in any of claims 8-21 or wherein the substance is a compound according to any of claims 22-25.

1/24

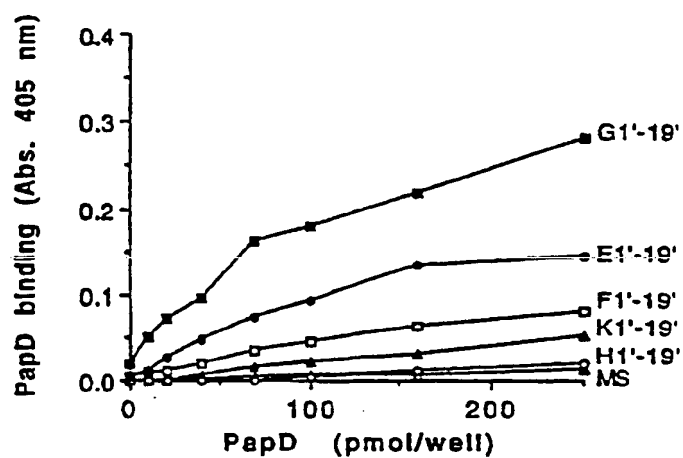


Fig. 1A

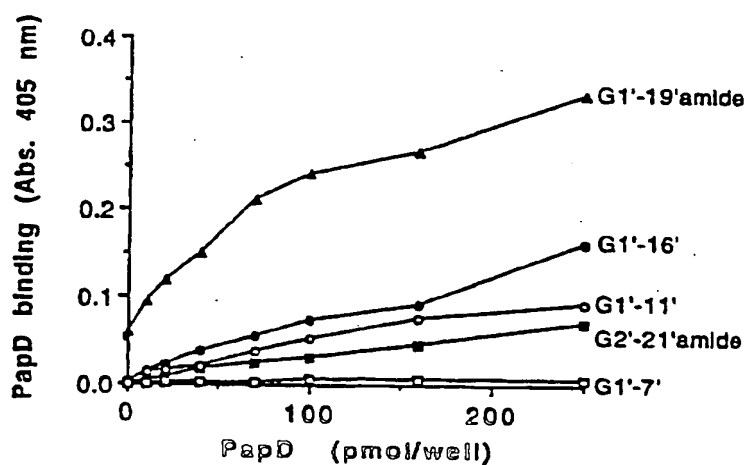


Fig. 1B

2/24

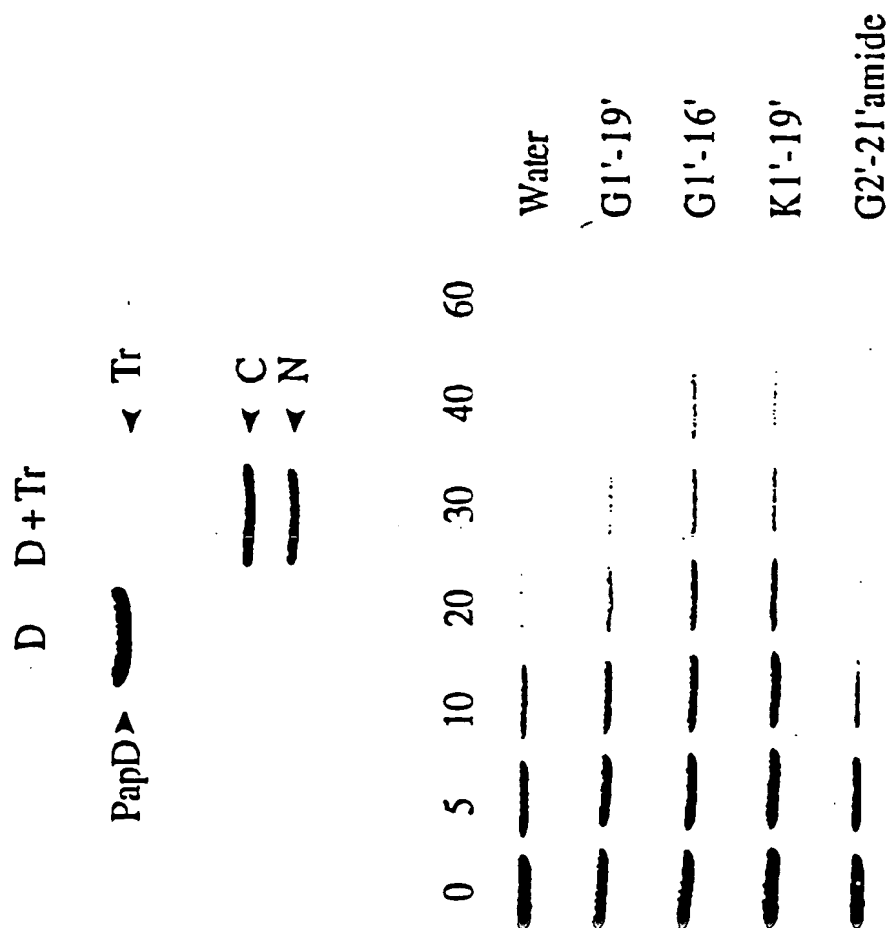


Fig. 2A

3/24

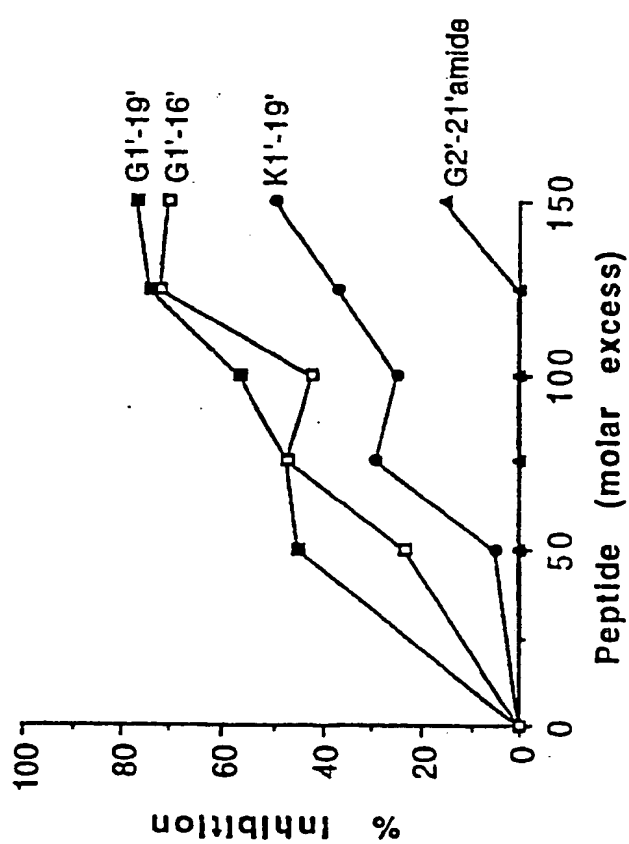
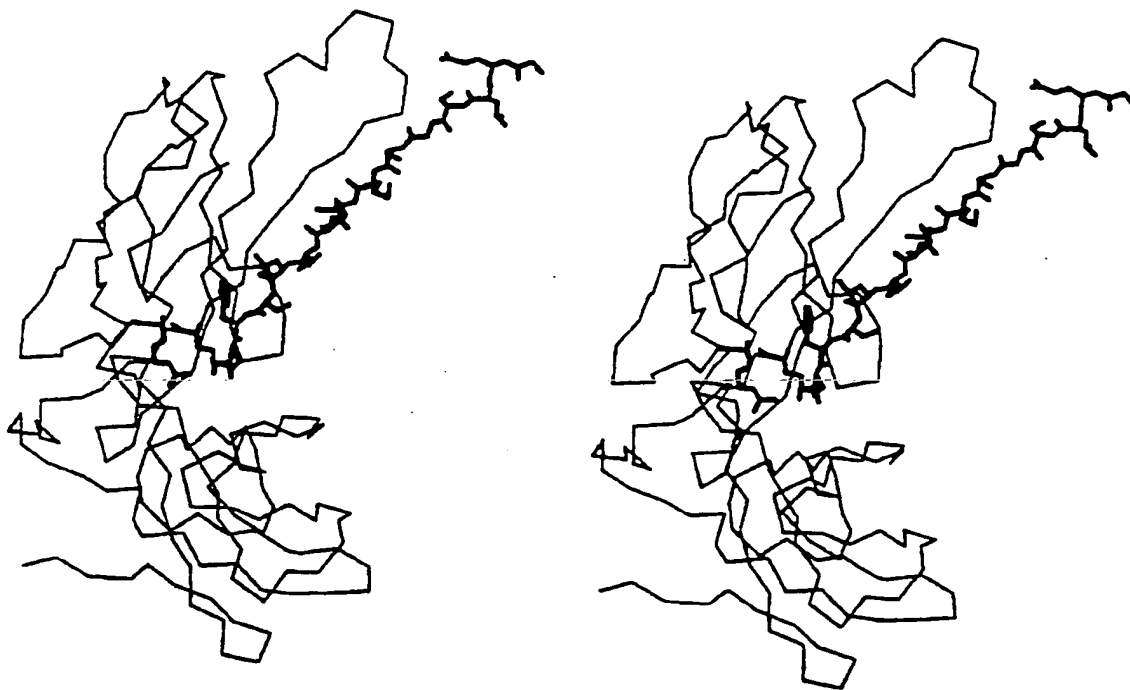


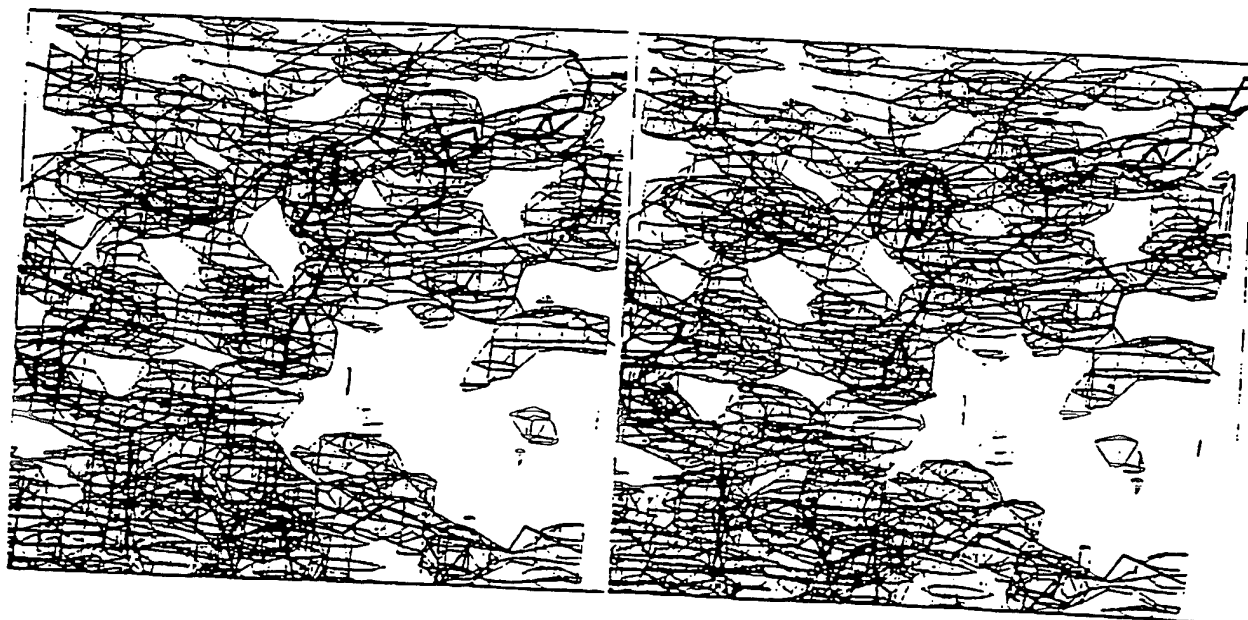
Fig. 2B

**Fig. 3**

4/24



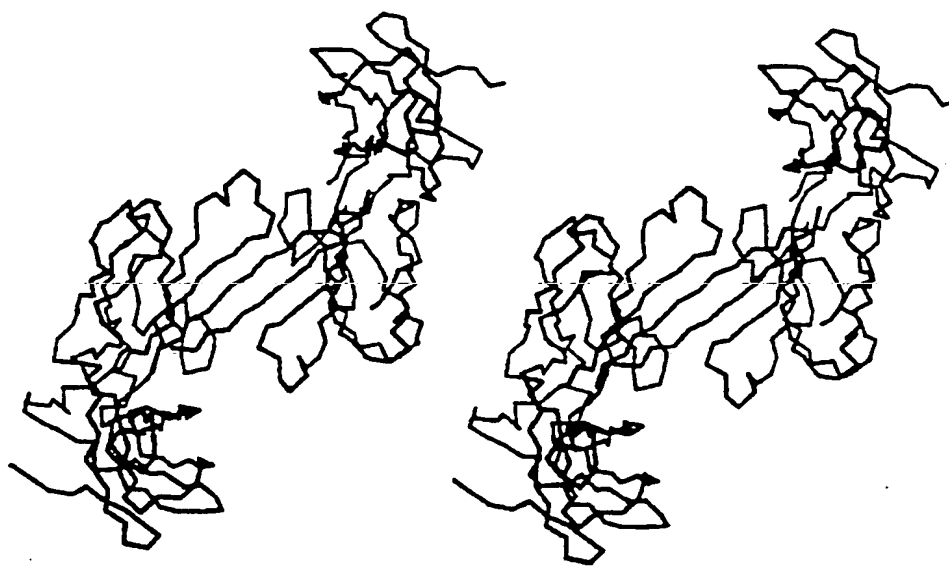
**Fig. 4**



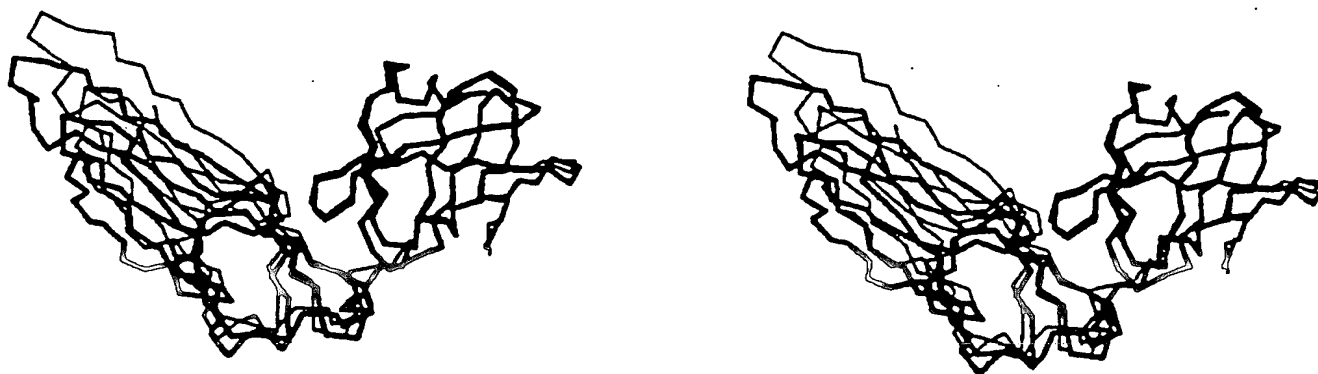


**Fig. 5**

**5/24**



**Fig. 6**



6/24

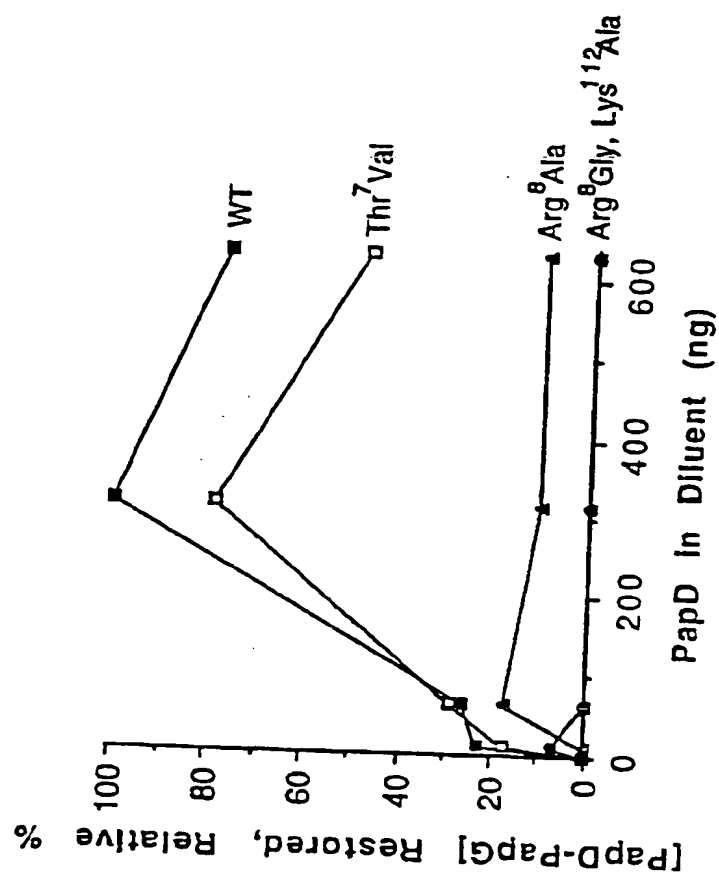


Fig. 7

7/24

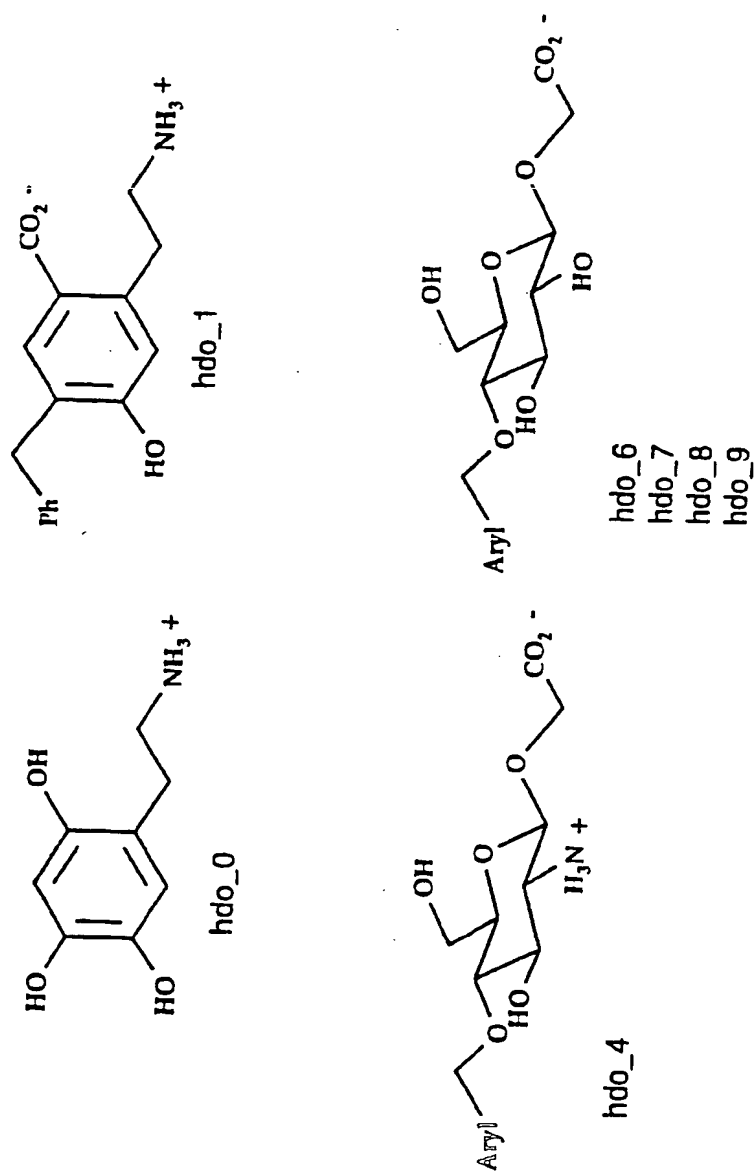


Fig. 8

8/24

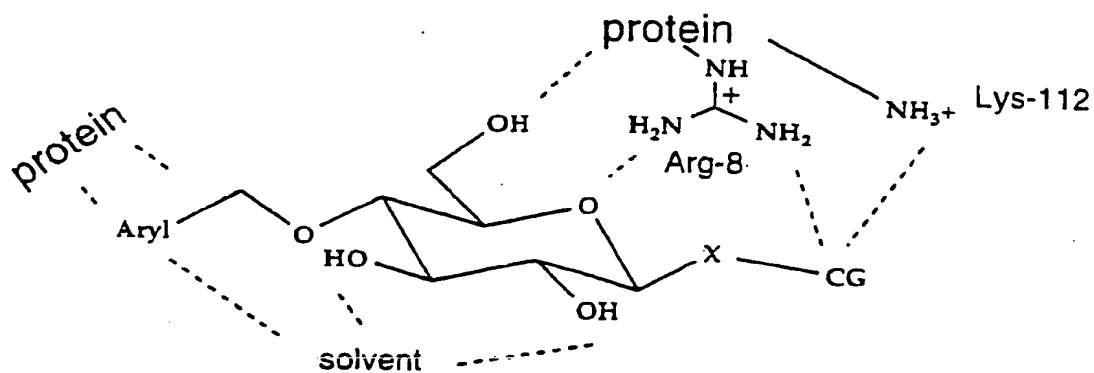


Fig. 9A

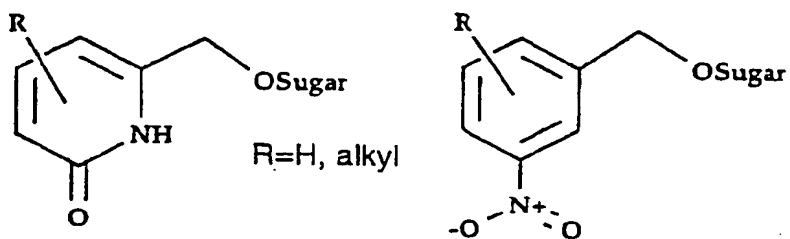


Fig. 9B

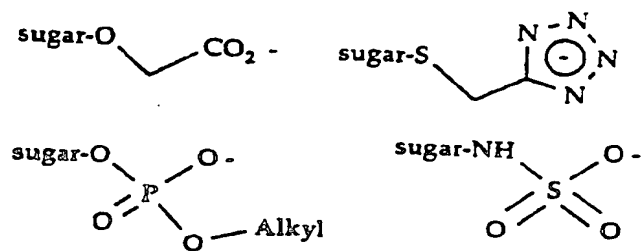


Fig. 9C

9/24

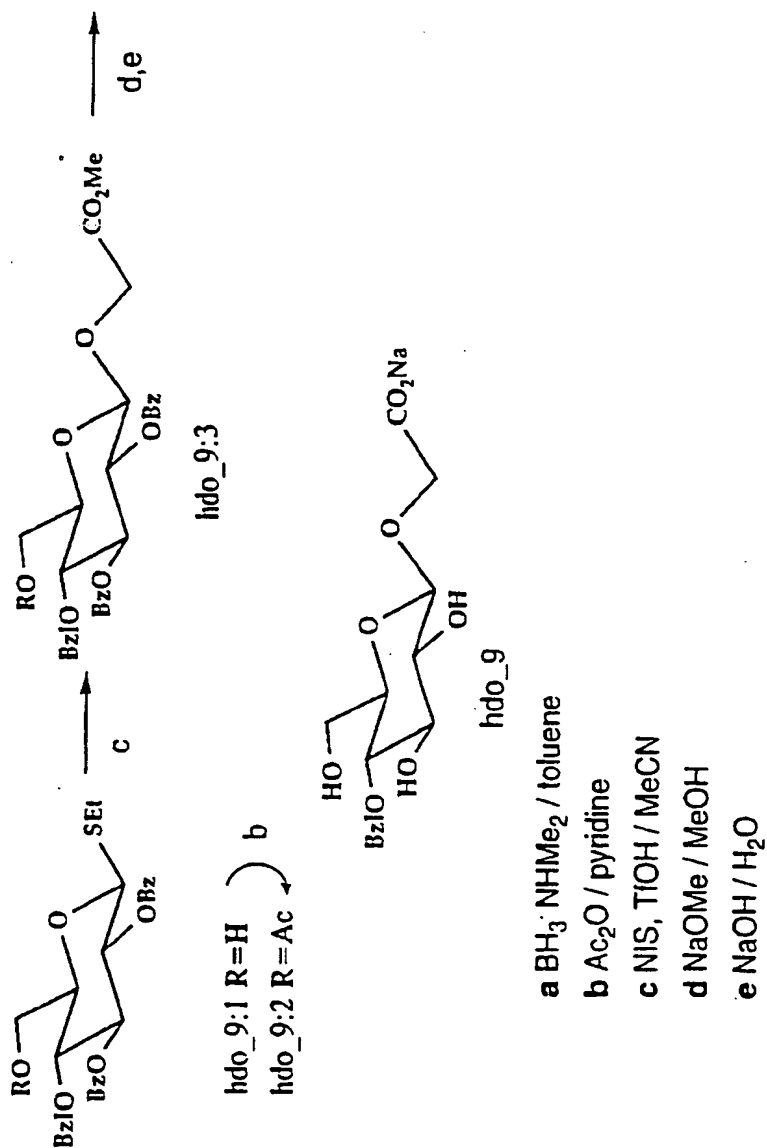


Fig. 10



11/24

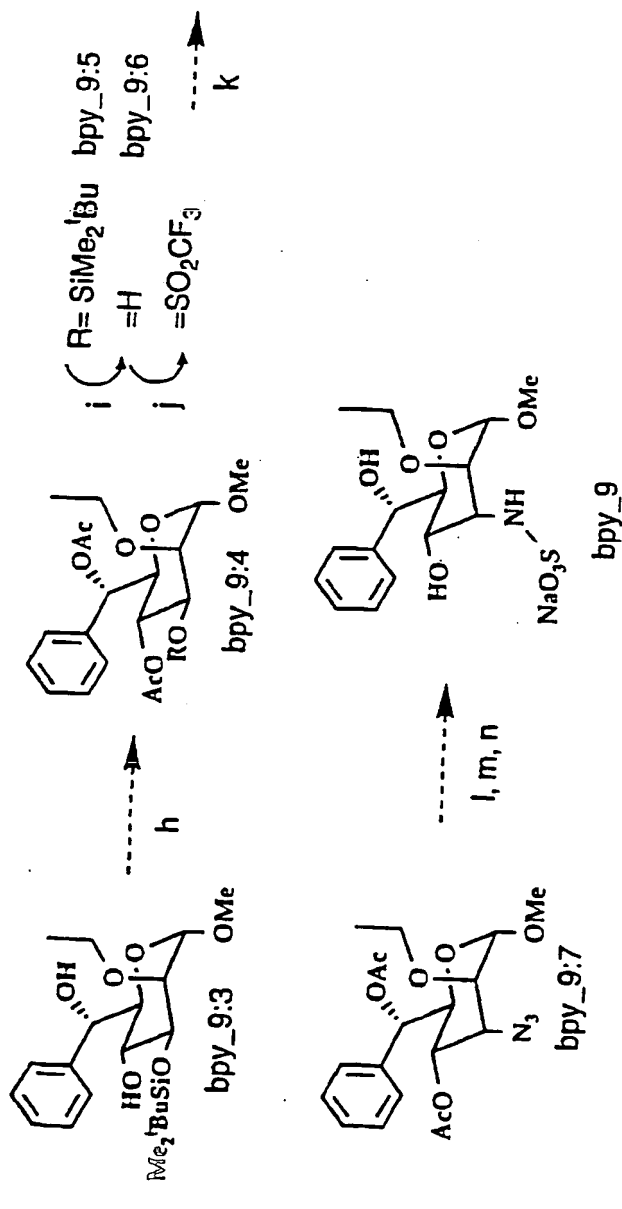


Fig. 12

12/24

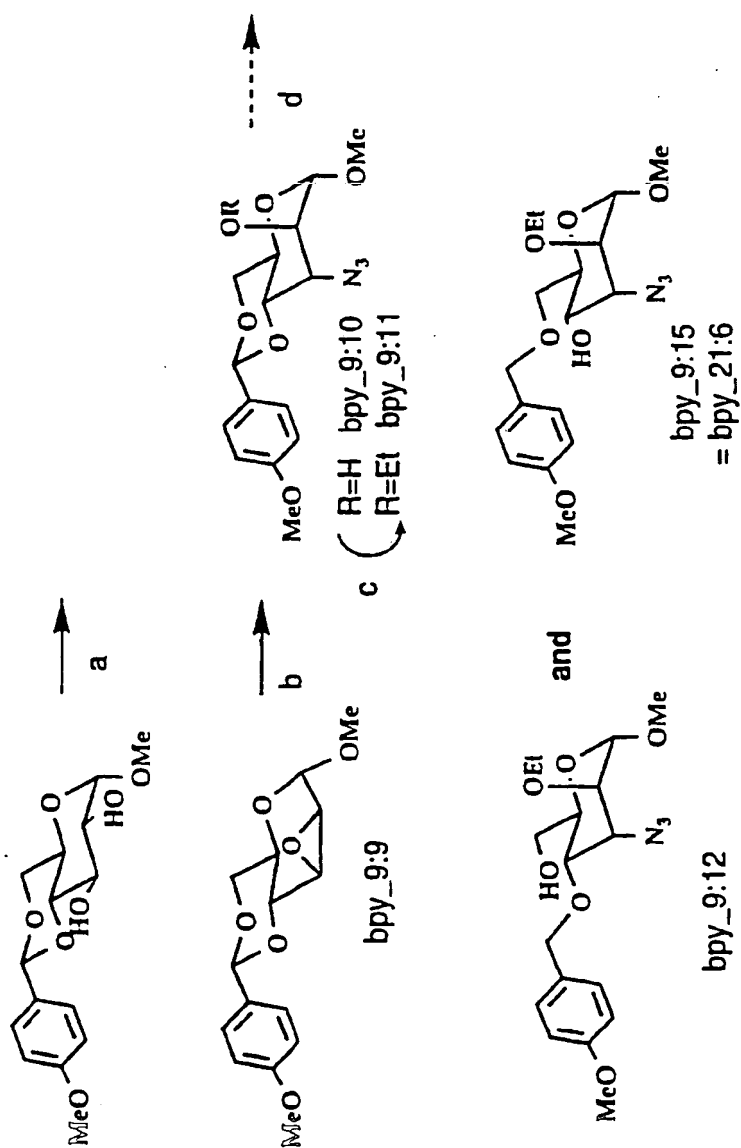
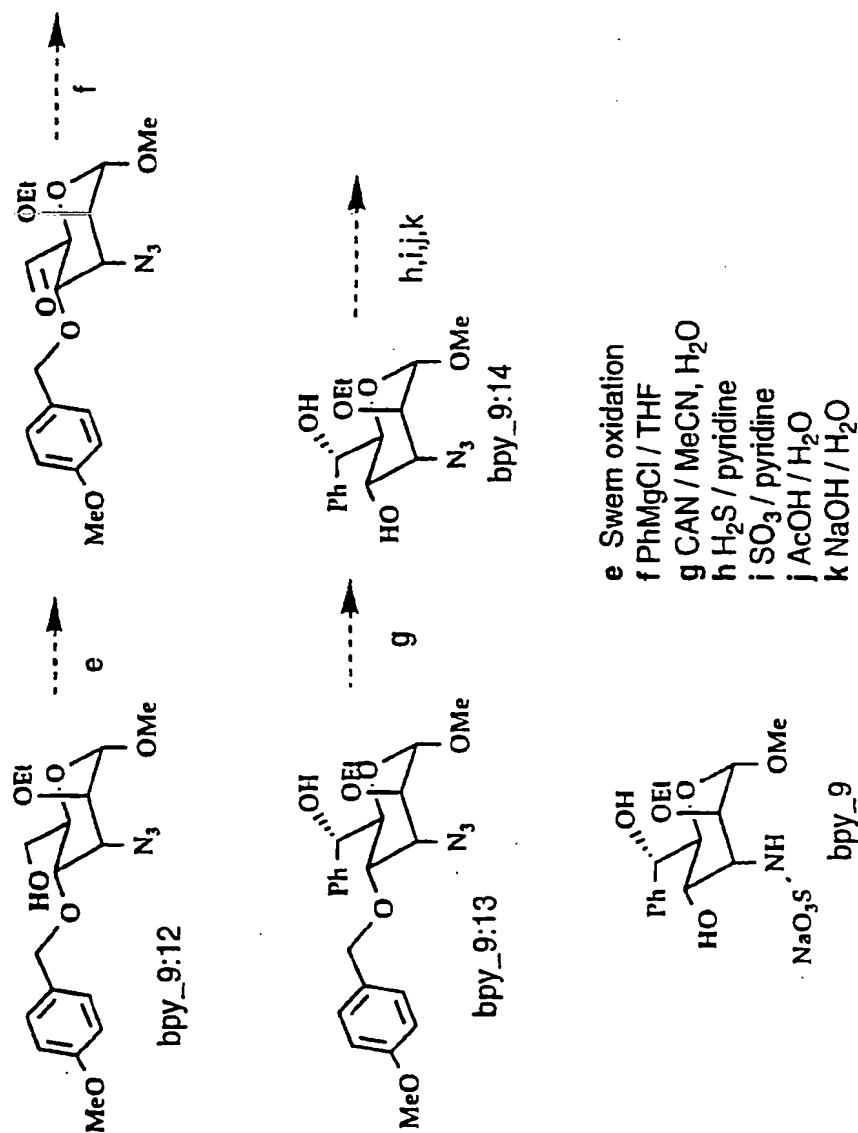


Fig. 13

a Imidazole tosylamide NaH / THF, rt, 66%  
 b NaN<sub>3</sub>, NH<sub>4</sub>Cl / EtOH, H<sub>2</sub>O, reflux, 72%  
 c EtI, NaH / DMF or alkyl halide, NaH / THF  
 d NaCNBH<sub>3</sub>, Me<sub>3</sub>SiCl / MeCN

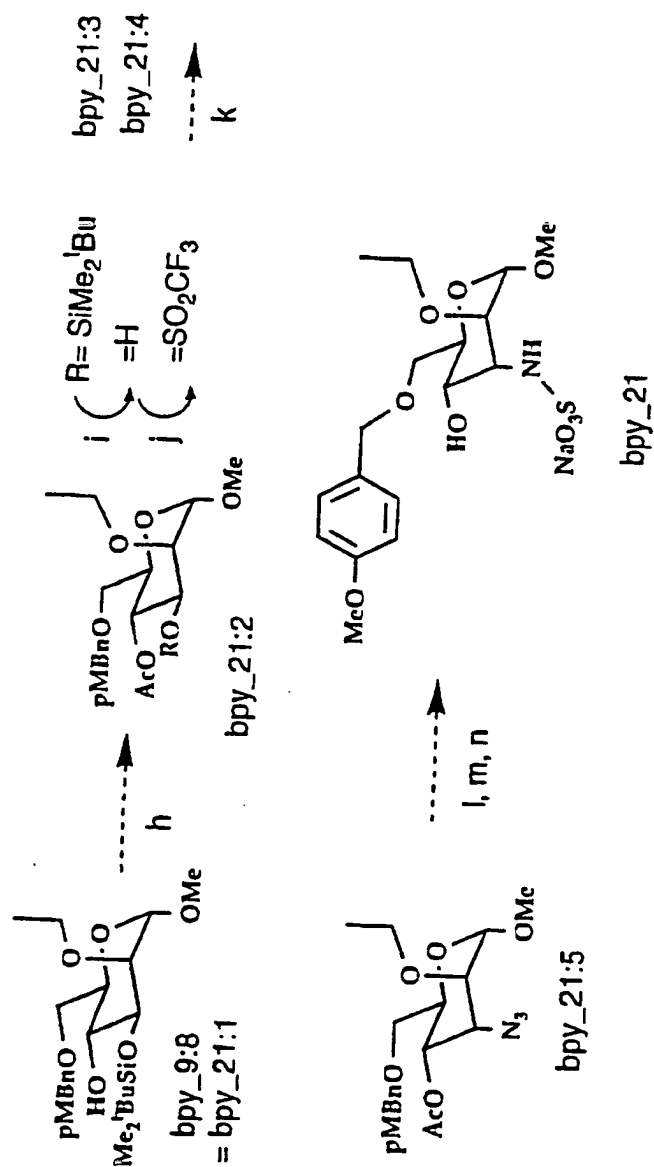


13/24



**Fig. 14**

14/24



- h Ac<sub>2</sub>O / pyridine  
 i NBu<sub>4</sub>F·3H<sub>2</sub>O / MeCN  
 j (CF<sub>3</sub>SO<sub>2</sub>)<sub>2</sub>O / CH<sub>2</sub>Cl<sub>2</sub>, pyridine  
 k NaN<sub>3</sub> / DMF  
 l H<sub>2</sub>S / pyridine  
 m SO<sub>3</sub> / pyridine  
 n NaOMe / MeOH

Fig. 15

15/24

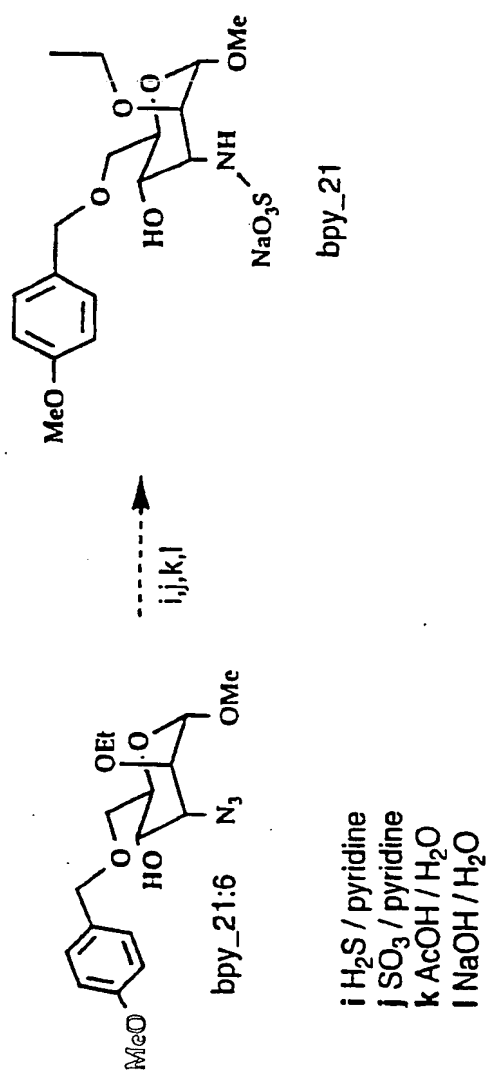


Fig. 16

16/24

Length series

(Data from 4 experiments)

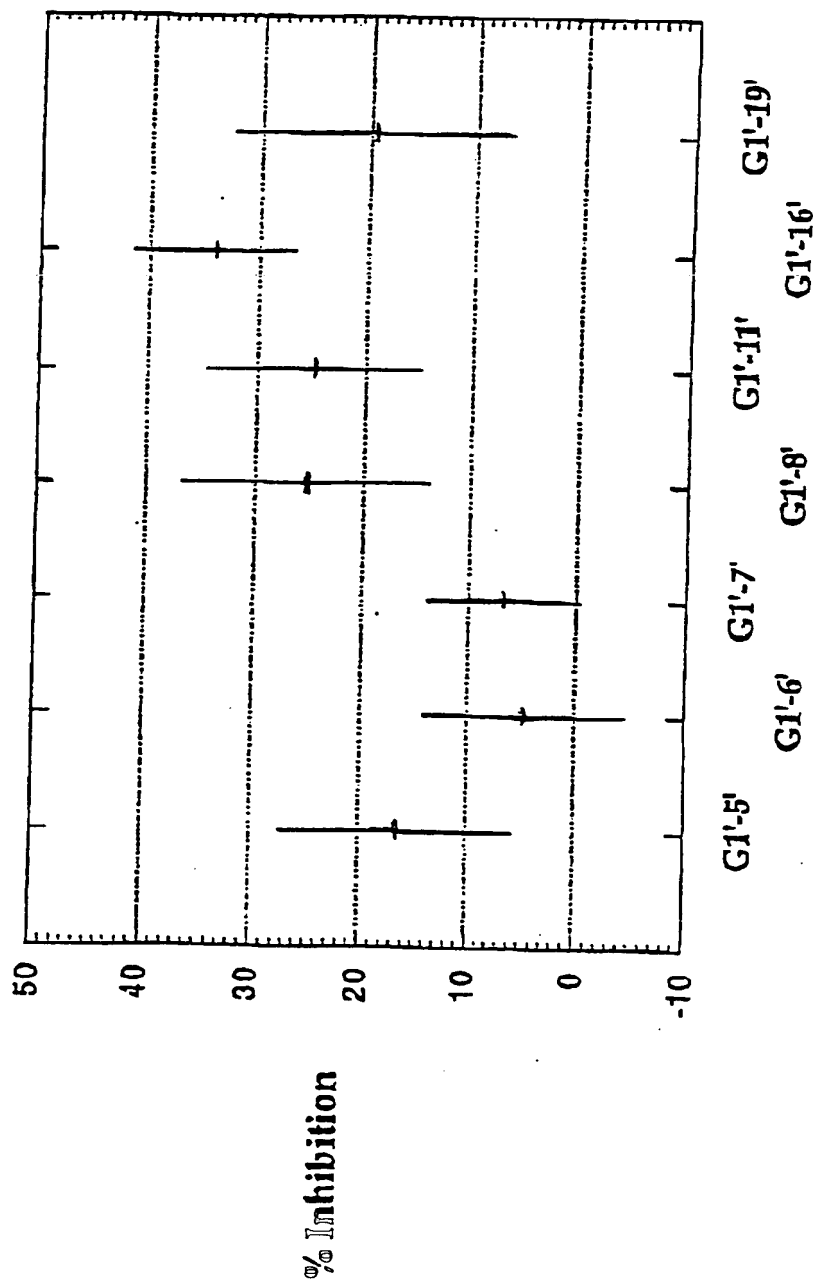


Fig. 17

17/24

Replacement series  
(Data from 7 experiments)

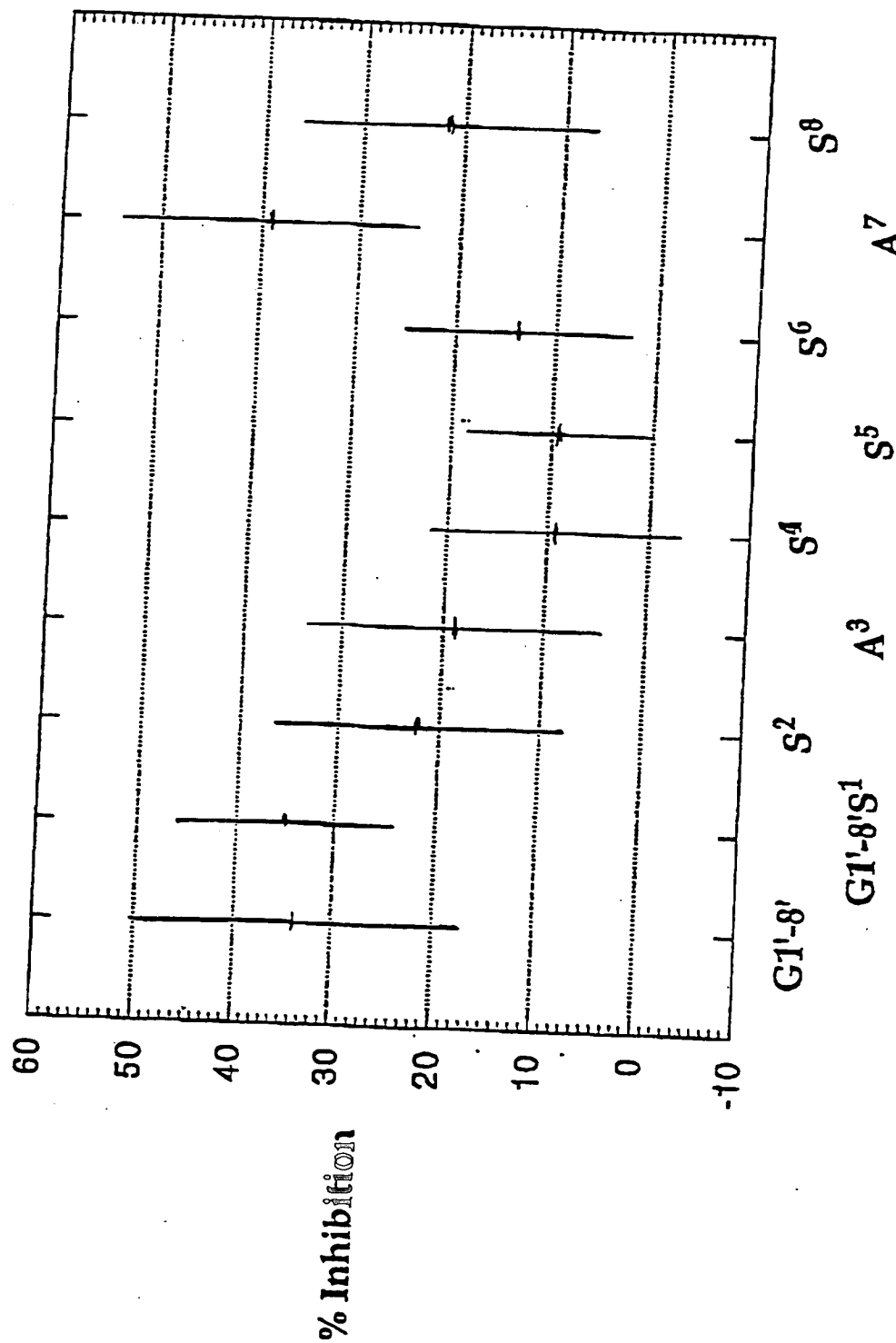


Fig. 18

18/24

Deletion series

(Data from 6/7 experiments)

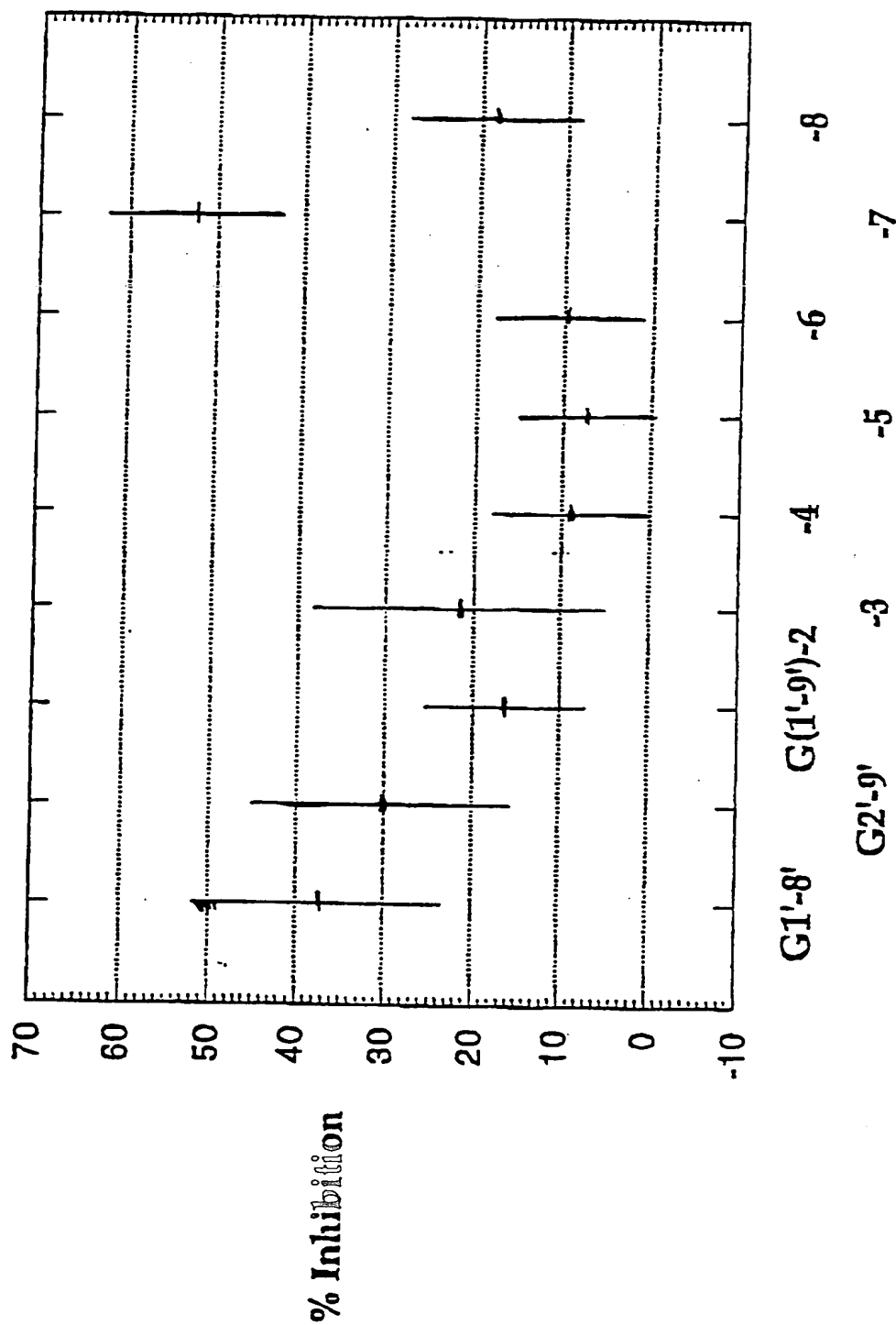


Fig. 19

19/24

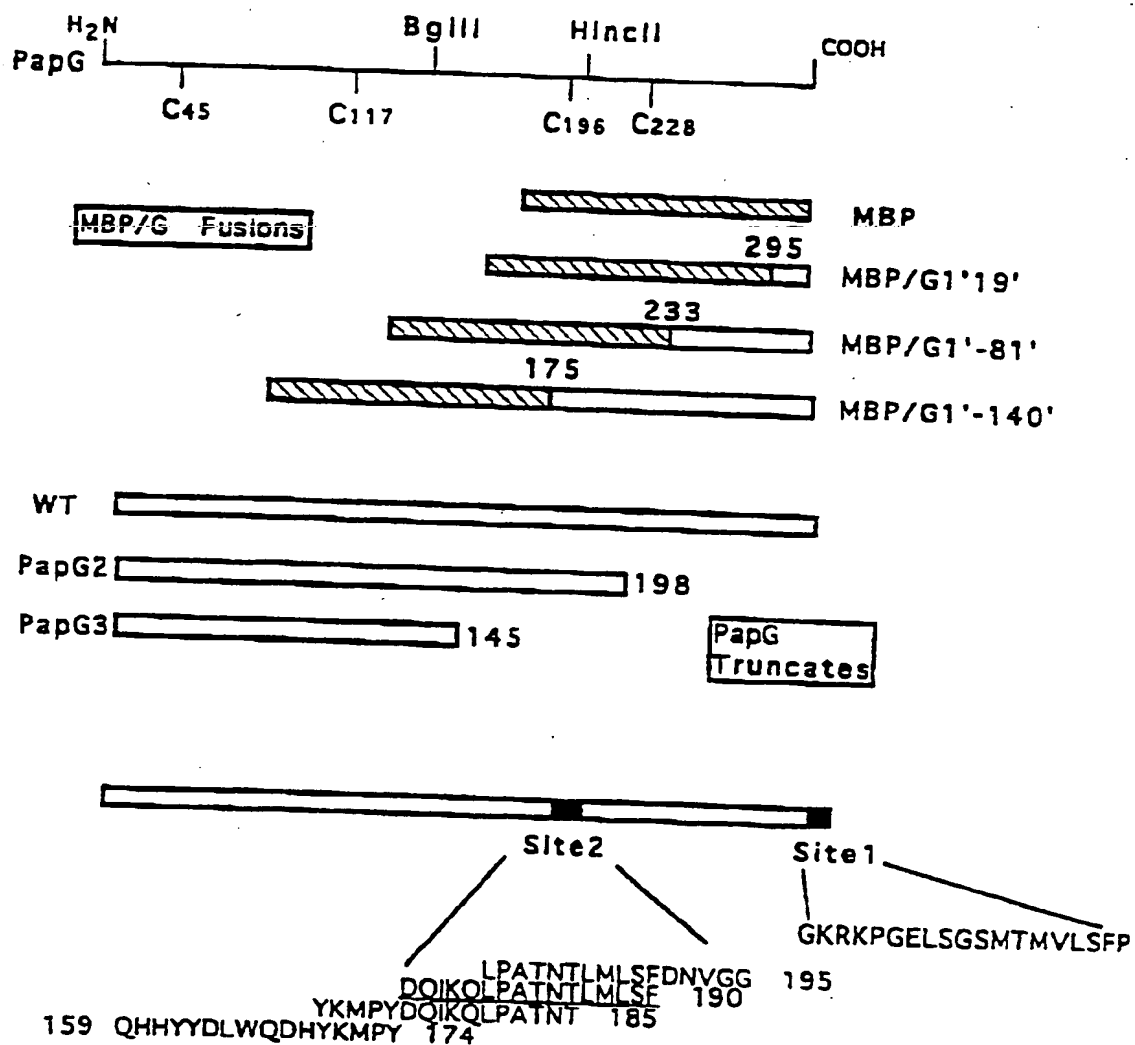


Fig. 20

20/24

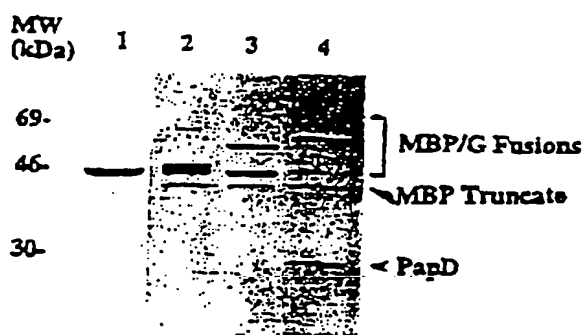


Fig. 21A

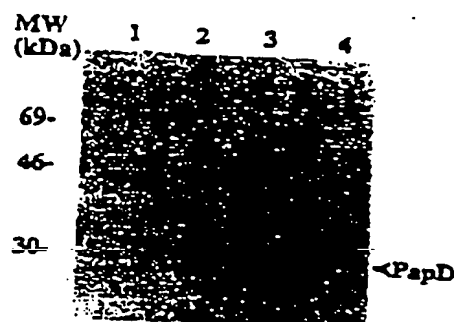


Fig. 21B

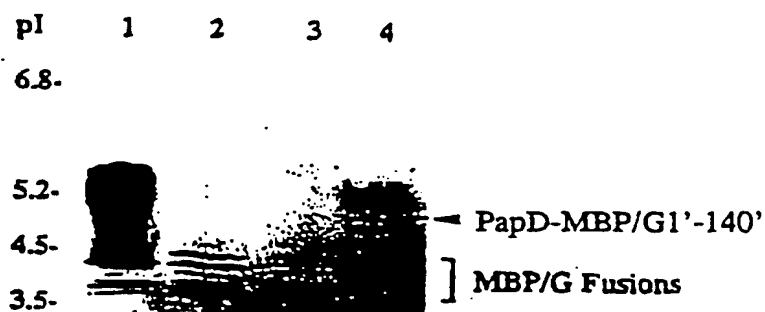


Fig. 21C

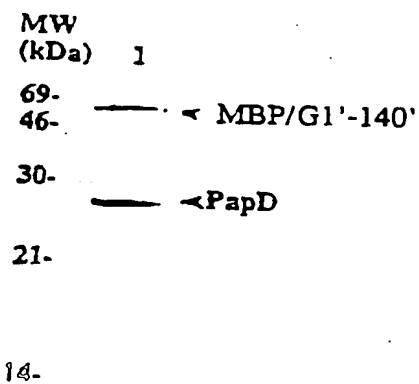


Fig. 21D



21/24

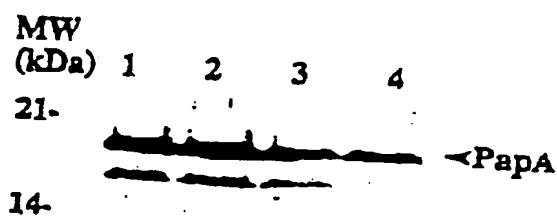


Fig. 22A

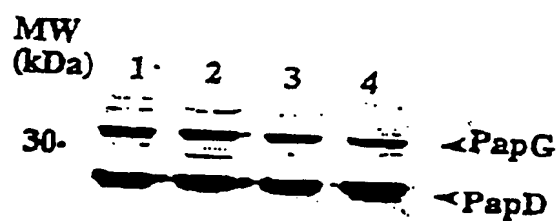


Fig. 22B

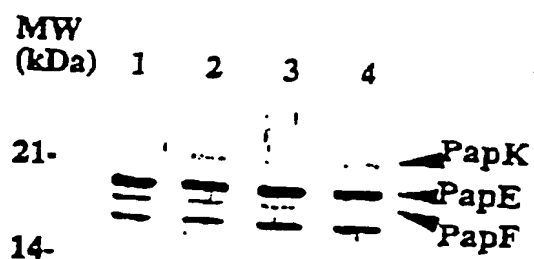


Fig. 22C

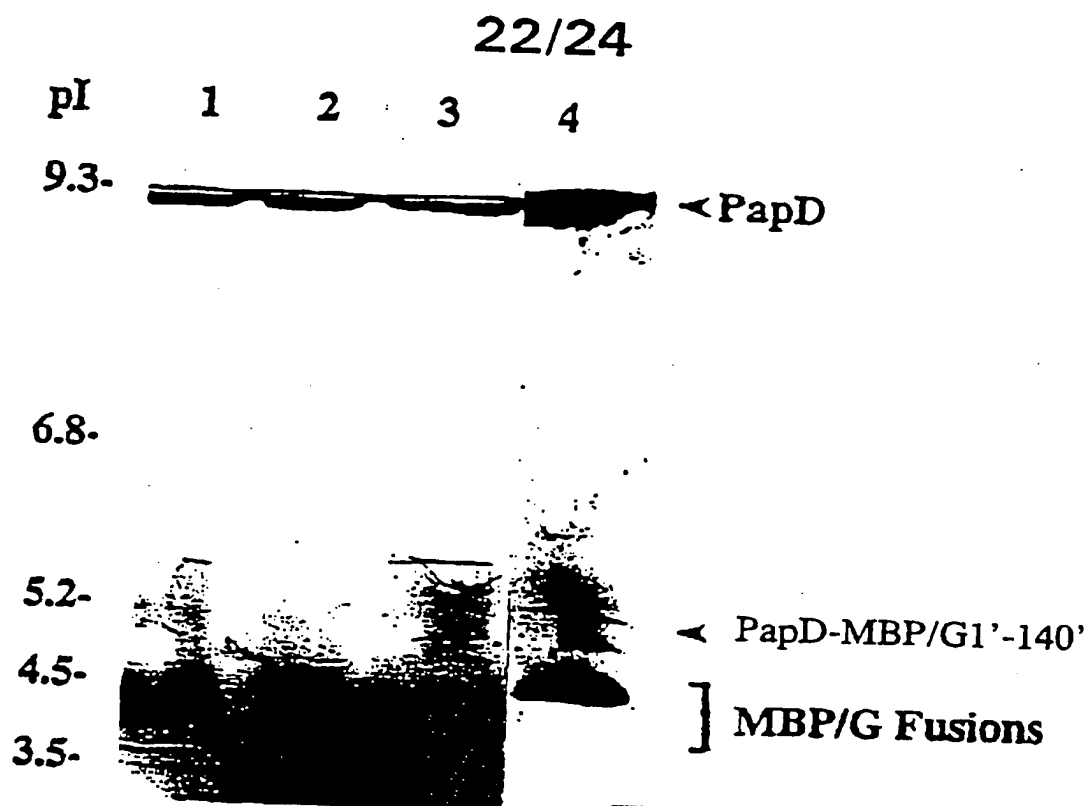


Fig. 23A

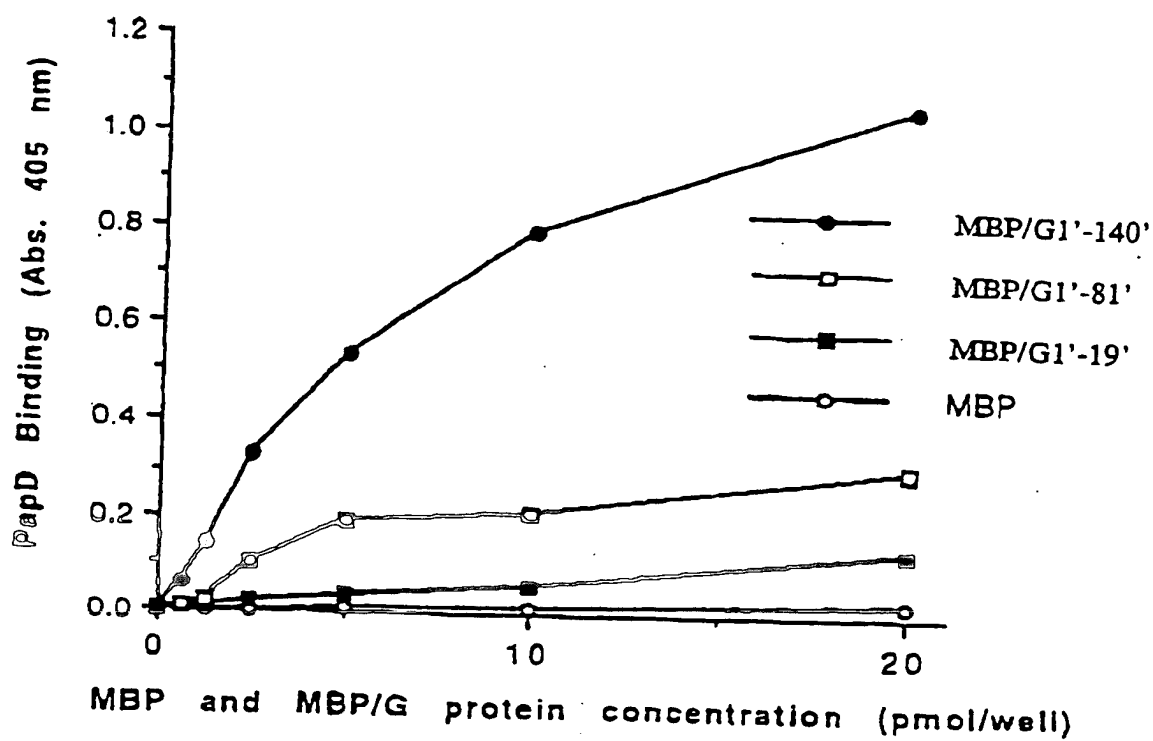


Fig. 23B

23/24

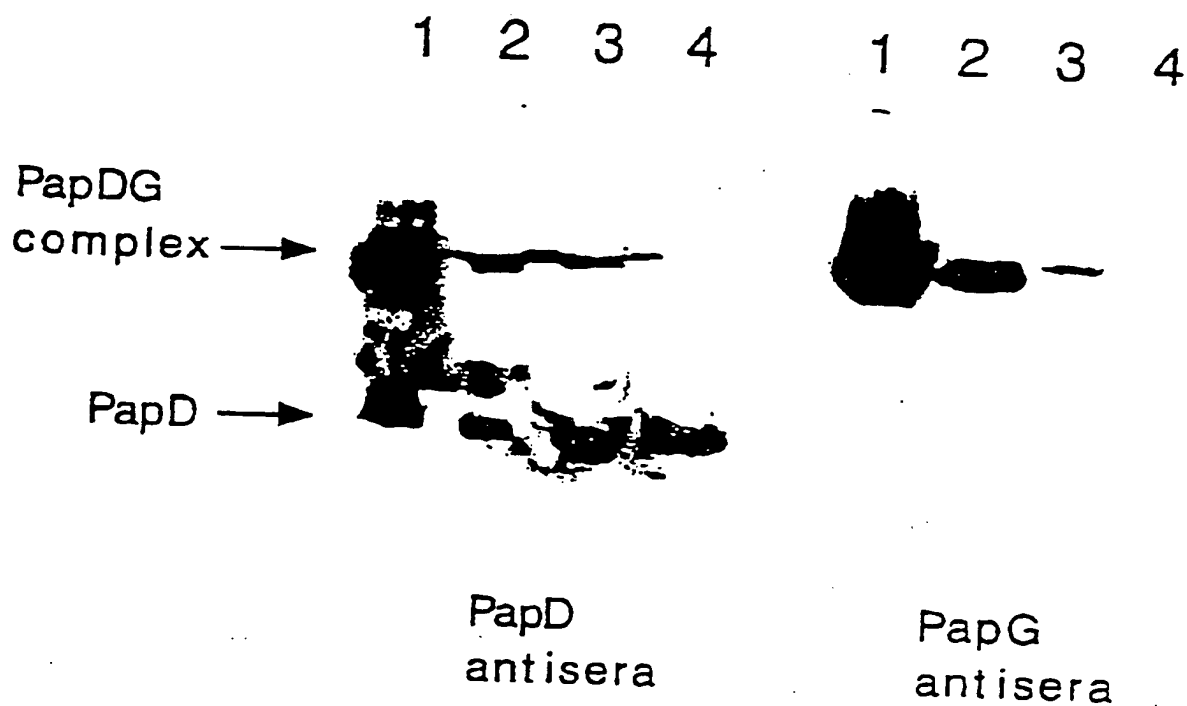


Fig. 24

24/24

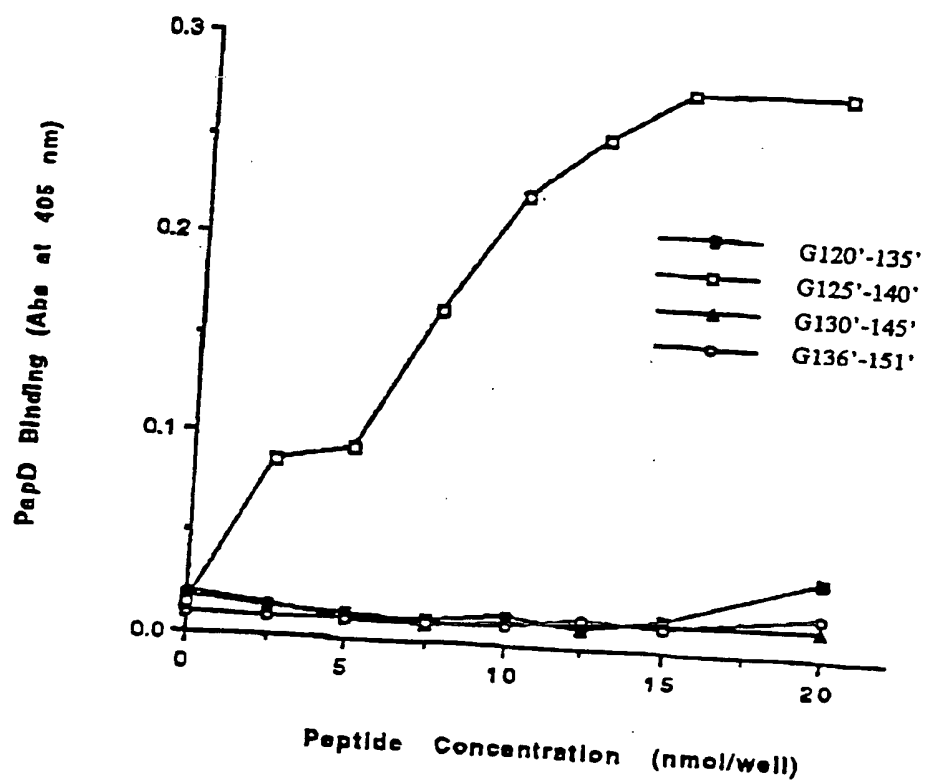


Fig. 25